Assessing the impact of mismodeled gaps in LISA data analysis

Ollie Burke (L2IT, Glasgow U.)

Sylvain Marsat (L2IT, Toulouse)

Jonathan Gair (AEI, Potsdam, DE)

Michael Katz (NASA Marshall SFC, USA)

[Burke&al arXiv:2502.17426, accepted in PRD]

Gaps in LISA data

- Gaps are missing data periods; can be **planned** and **unplanned**
- The stochastic noise process could be **coherent** (masked data) or **incoherent** (instrument reset)
- Masking out data is also a crude way of addressing glitches

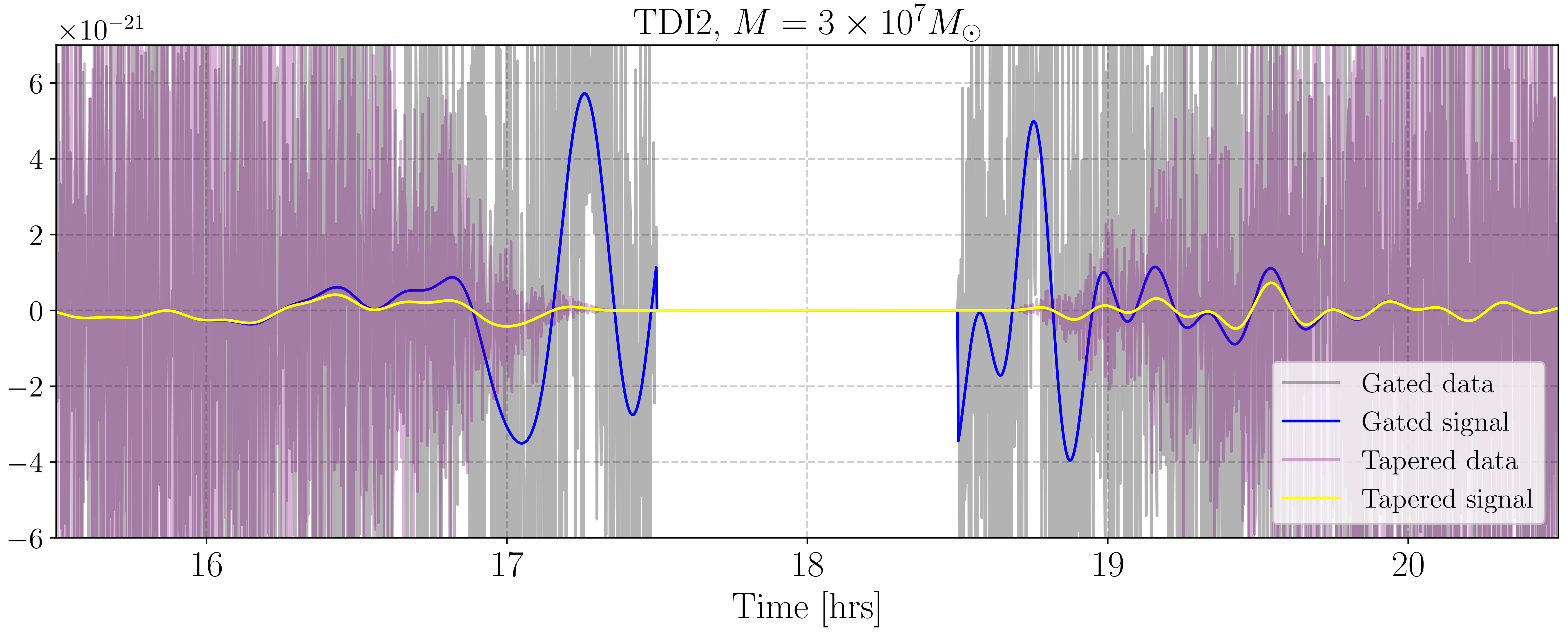
Gap type	Frequency	Duration	Total loss (hr/yr)
Antenna repointing	every 2 weeks	3.3h	1%
PAAM angle adjust	3 per day	100s	0.3%
TM stray pot. est.	2/yr	1 day	0.56%
TTL coupling est.	4/yr	2 days	2.22%
Unplanned: platform	3/yr	2.5 days	2%
Unplanned: payload	4/yr	2.75 days	3%
Unplanned: micro-meteorites	30/yr	1 day	8%

[A. Petiteau, FMT]

Gaps in LISA data: different approaches

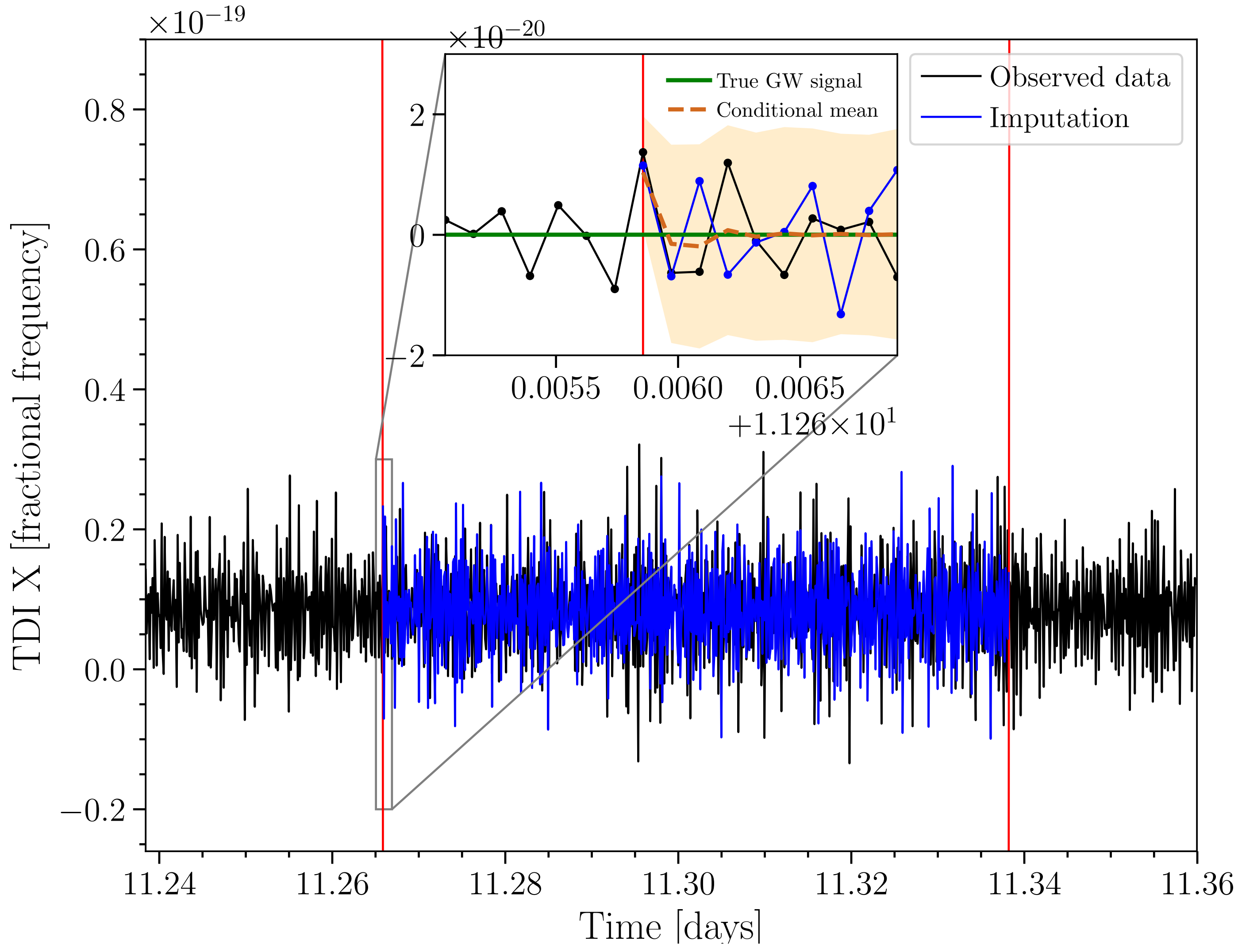
- **Windowing** the data around gaps, having well-behaved Fourier-domain signals

  [Dey&al arXiv:2104.12646]



- **Data augmentation** (or imputation): the data inside the gaps is inferred in a Gibbs sampling between GW params and missing data

  [Baghi&al arXiv:1907.04747]



- **Modelling of the full covariance:** direct marginalization over missing data, useful as a test bed for our assumptions — limited by the size of the $N \times N$ covariance matrix — **This work**

The basics: Stationary Gaussian process

- Assumption: noise as **Gaussian process** described by its covariance $C(t, t') = \langle n(t) n(t') \rangle$

- Assumption: underlying noise (before introducing gaps) **Stationary**, with autocorrelation depending on lag only

$$C(t, t') = C(0, t' - t) \equiv C(t' - t)$$

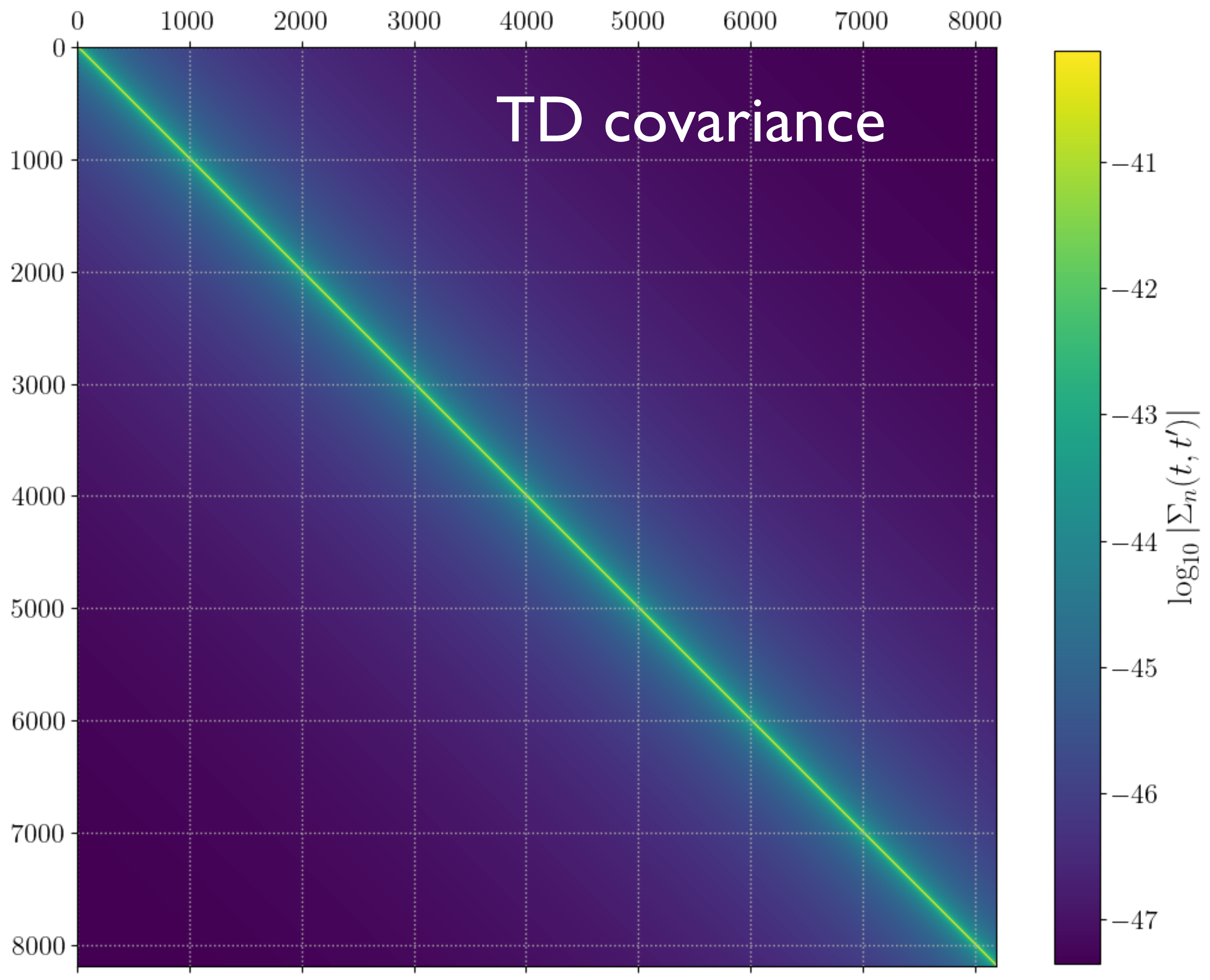
- In the Fourier domain, this leads to independence:

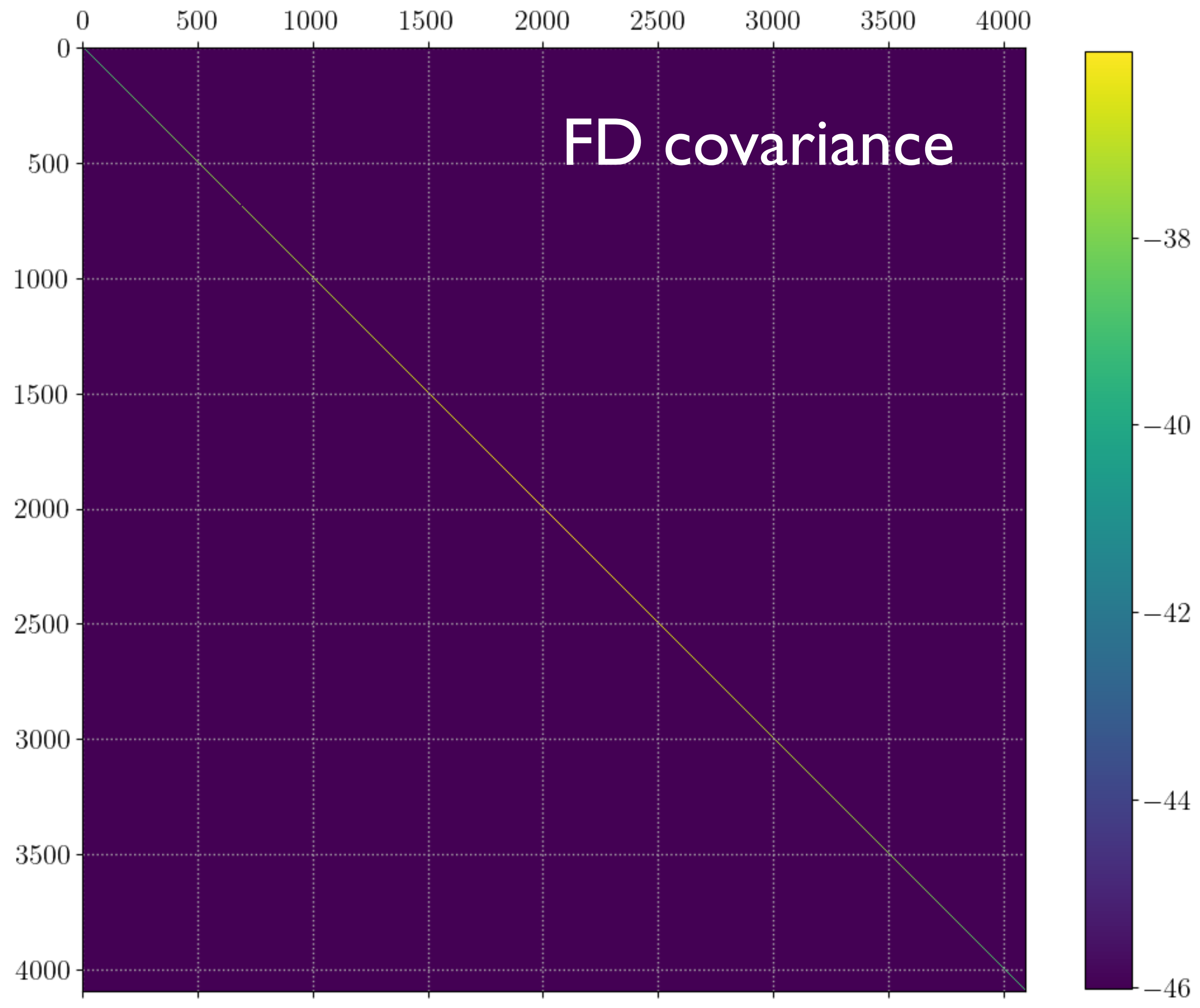
$$\langle \tilde{n}(f) \tilde{n}^*(f') \rangle = \frac{1}{2} S_n(f) \delta(f - f')$$

- **Whittle likelihood**:

$$\ln \mathcal{L}(\theta) = -\frac{1}{2}(h(\theta) - d | h(\theta) - d) \quad (a|b) = 4\mathrm{Re} \int \frac{df}{S_n(f)} \tilde{a}^*(f) \tilde{b}(f)$$

- Discrete time and Fourier domain: $\langle N N^T \rangle = \Sigma$ $\quad \langle \tilde{N} \tilde{N}^\dagger \rangle = \tilde{\Sigma}$ The Fourier-domain covariance matrix is diagonal: $N \times N \rightarrow N$!



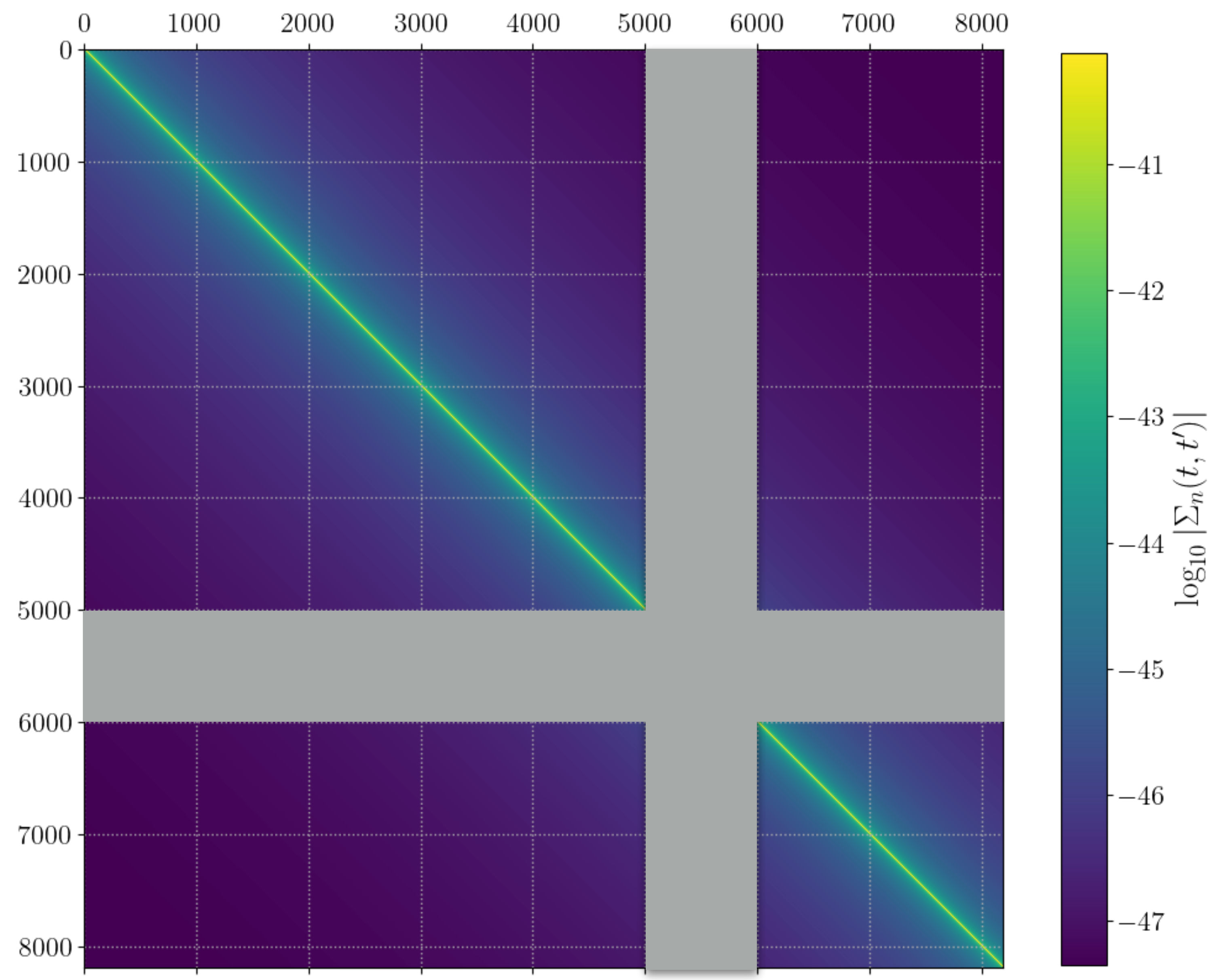


Gaps will break the stationarity assumption

[Stationarity $\Longrightarrow$ **Toeplitz**
Circulant $\Longrightarrow$ diagonal covariance FD]

Marginalizing out the missing data: Time Domain

Direct marginalization:



Truncation of the Gaussian covariance:

$$X = H(\theta) - D = -N \qquad X = (X_1, X_2)$$

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12}^T & \Sigma_{22} \end{pmatrix}$$

$$\mathcal{L}_{\text{gap}} = p(N_1) = \int dN_2 \, p(N_1, N_2)$$

$$\ln \mathcal{L}_{\text{gap}} = -\frac{1}{2} X_1^T \Sigma_{11}^{-1} X_1$$

Gap seen as a gated process:

For an invertible modulation W:

$$X \to WX$$

$$X^T \Sigma^{-1} X = (WX)^T (W\Sigma W)^{-1} (WX)$$

For an gating function W, non-invertible: $W = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & 0 \end{pmatrix}$

Moore-Penrose **pseudo-inverse** (uniquely defined):

$$A^+ A A^+ = A^+$$
$$A A^+ A = A$$
$$(AA^+)^\dagger = AA^+$$
$$(A^+ A)^\dagger = A^+ A$$

$$(W\Sigma W)^+ = \begin{pmatrix} \Sigma_{11}^{-1} & 0 \\ 0 & 0 \end{pmatrix}$$

Equivalence with the marginalized likelihood:

$$X^T (W\Sigma W)^+ X = X_1^T \Sigma_{11}^{-1} X_1$$

Likelihood in time or Fourier domain:

$$\ln \mathcal{L}_{\text{gap}} = -\frac{1}{2} (WX)^T (W\Sigma W)^+ (WX) = -\frac{1}{2} \left(\widetilde{WX}\right)^\dagger \left(\tilde{W}\tilde{\Sigma}\tilde{W}\right)^+ \left(\widetilde{WX}\right)$$

Effect of noise on posterior and Cutler-Vallisneri biases

Linearized signal approximation:

$$H(\theta) \simeq H(\theta_0) + \Delta\theta^i \partial_i H$$

Likelihood: $\ln \mathcal{L} \simeq -\frac{1}{2}\Gamma_{ij}\Delta\theta^i\Delta\theta^j$

Fisher information matrix and covariance:

$$\Gamma_{ij} = (\partial_i H | \partial_j H)_\Sigma \qquad \Gamma^{-1}_{ij}$$

Cutler-Vallisneri bias due to noise realization:

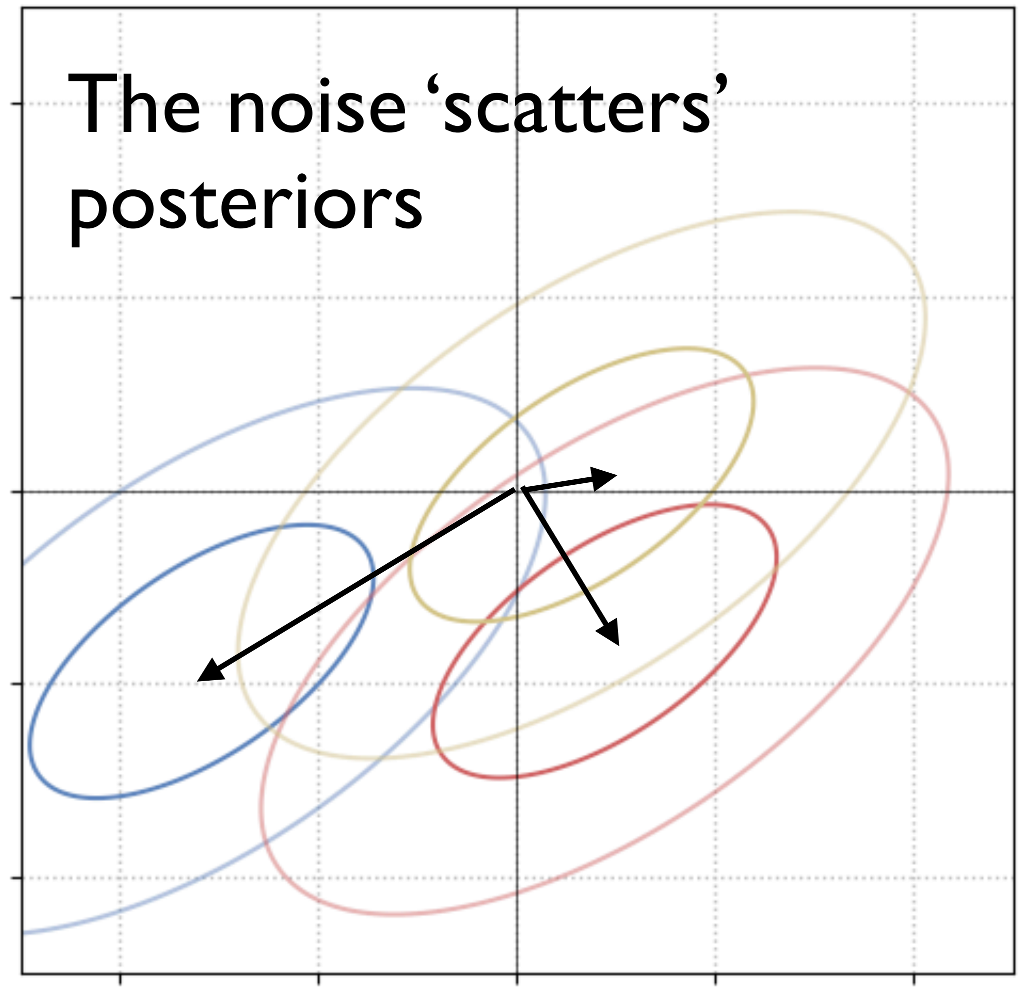
$$D = H(\theta_0) + N$$

CV bias and variance: $\Delta\theta^i_{\rm bf} = \Gamma^{-1}_{ij}(\partial_j H | N)_\Sigma$

$$\langle \Delta\theta^i_{\rm bf} \rangle = 0$$

$$\langle N N^T \rangle = \Sigma \longrightarrow \langle \Delta\theta^i_{\rm bf}\Delta\theta^j_{\rm bf} \rangle = \Gamma^{-1}_{ij}$$

The CV bias variance matches the Fisher covariance



Introduce mismodelling in the covariance itself:

- Wrong covariance Σ' assuming stationarity (Whittle), while gating the signals
- Could also be used to introduce any misestimation of the PSD/covariance

Fisher matrix errors:

$$\Gamma'_{ij} = \partial_i (WH)^T (\Sigma')^{-1} \partial_j (WH)$$ Fisher matrix for the approx. likelihood

$$(\Gamma')^{-1}_{ij} \neq (\Gamma)^{-1}_{ij}$$ the size of Fisher errors is different

Covariance of noise-induced bias:

$\Delta\theta^i_{\rm bf'} = \Gamma'^{-1}_{ij}(W\partial_j H | WN)_{\Sigma'}$ $\quad \langle \Delta\theta^i_{\rm bf'} \rangle = 0$ displacement still zero-mean

$$\langle \Delta\theta^i_{\rm bf'}\Delta\theta^j_{\rm bf'} \rangle = \Gamma'^{-1}_{ik}\Gamma'^{-1}_{jl}\partial_k H^T W \Sigma'^{-1} W \Sigma W \Sigma'^{-1} W \partial_l H$$ differs from covariance

[Edy&al arXiv:2101.07743]

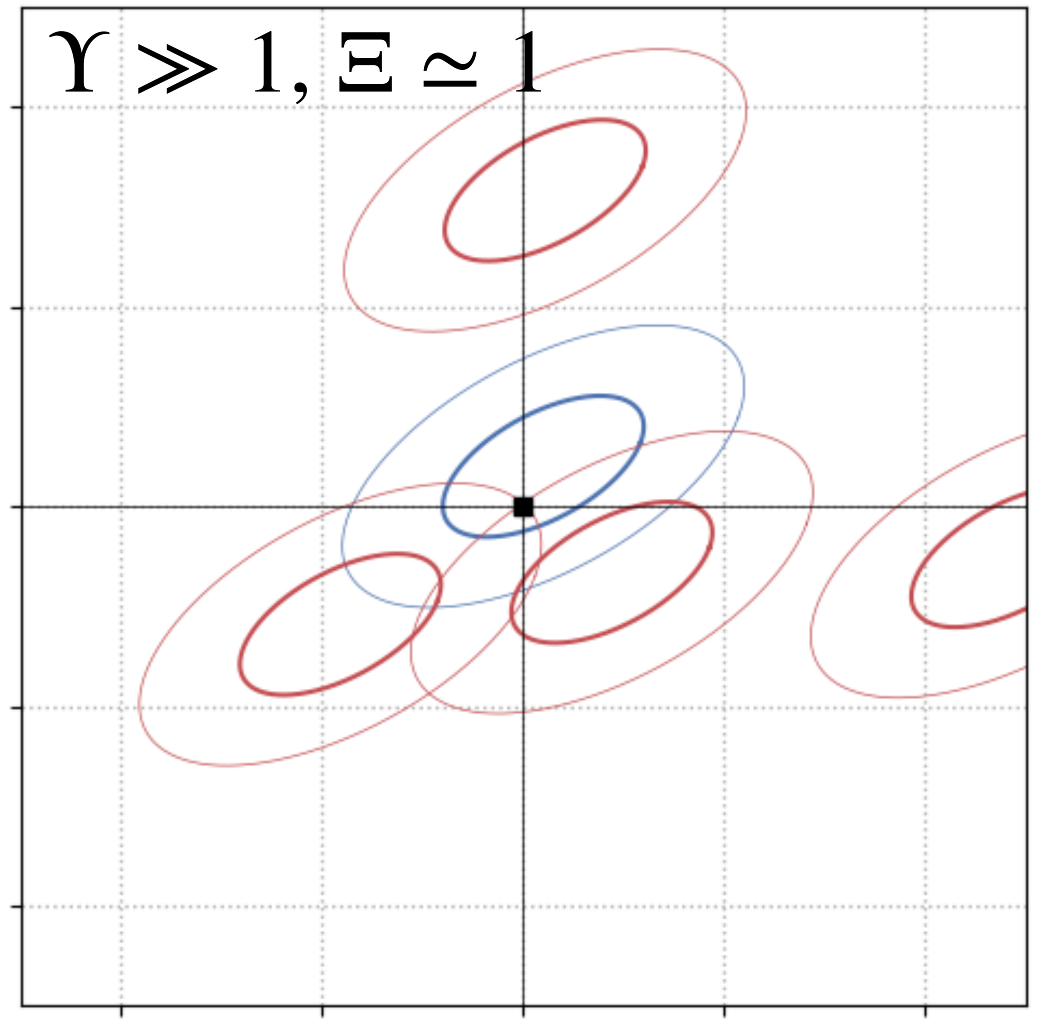
Mismodelling: error measures

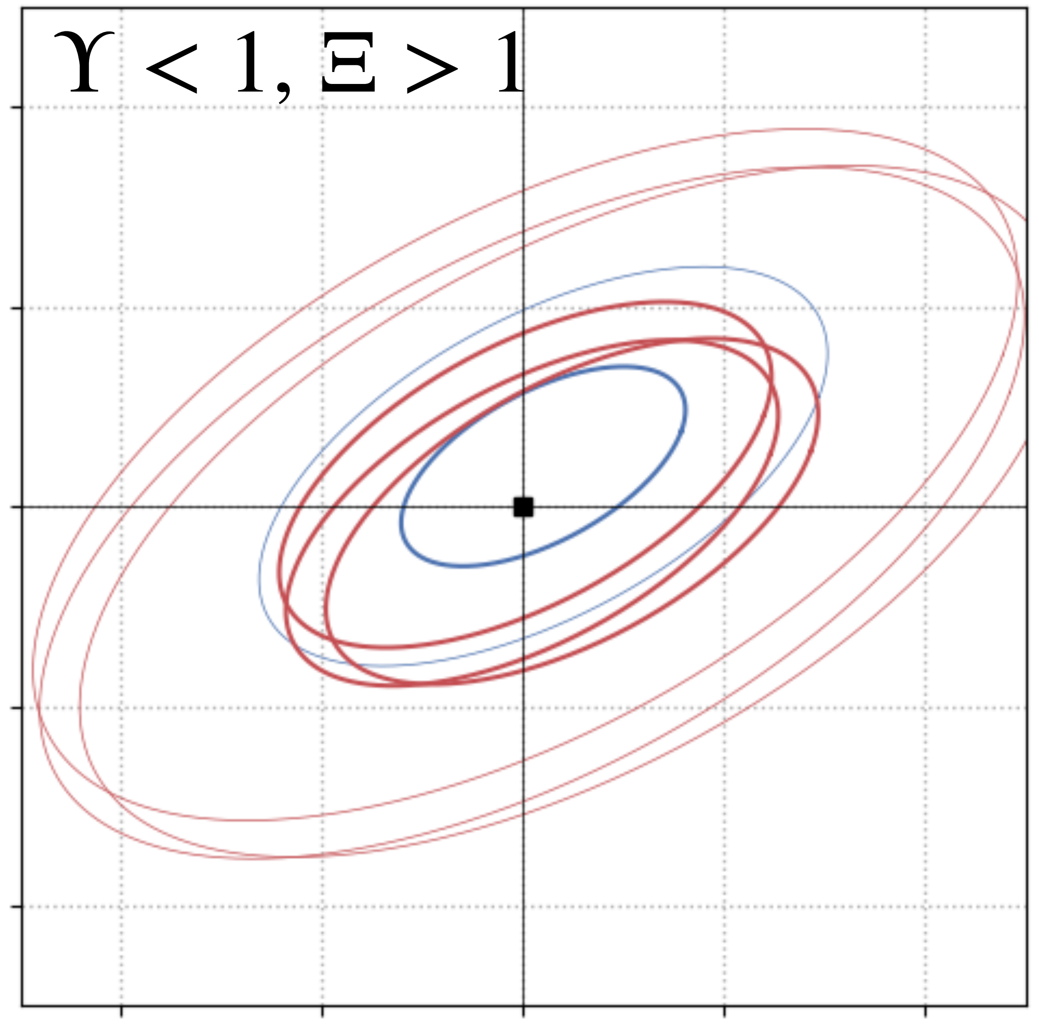
Noise mismodelling posterior scatter-to-width ratio :

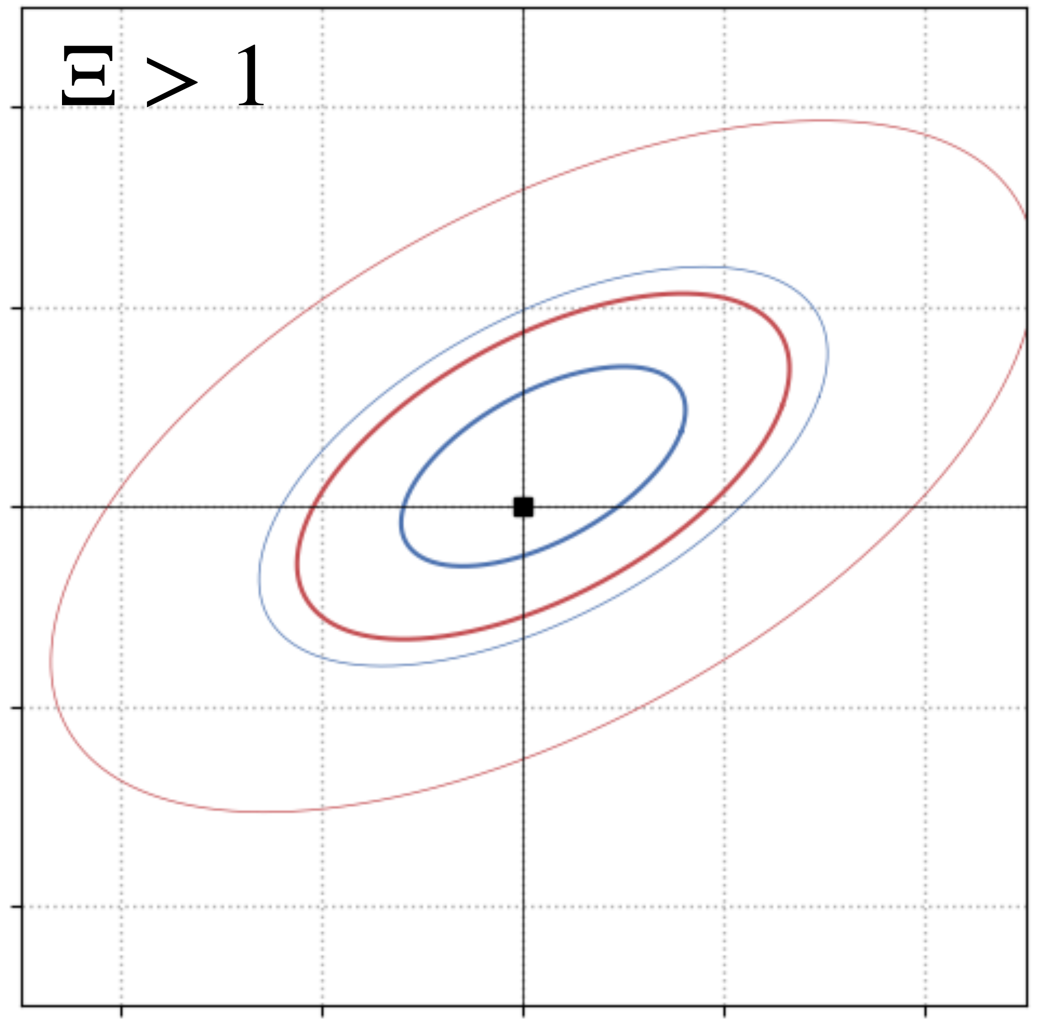
$$\Upsilon_a = \frac{\sqrt{\left\langle \left(\Delta\hat{\theta}^a_{\rm bf}\right)' \left(\Delta\hat{\theta}^a_{\rm bf}\right)' \right\rangle}}{\sqrt{(\mathbf{\Gamma}')^{-1}_{aa}}}$$

Noise mismodelling posterior width ratio :

$$\Xi_a = \frac{\sqrt{(\mathbf{\Gamma}')^{-1}_{aa}}}{\sqrt{(\mathbf{\Gamma})^{-1}_{aa}}}$$







- Υ measures statistical inconsistency: $\Upsilon > 1$ for too much scatter of the posteriors, $\Upsilon < 1$ for not enough
- Different from systematic bias (waveform errors): here displacement is random and always 0-mean
- Proxy for p-p plots
- Can be computed directly or from (any) simulated noise

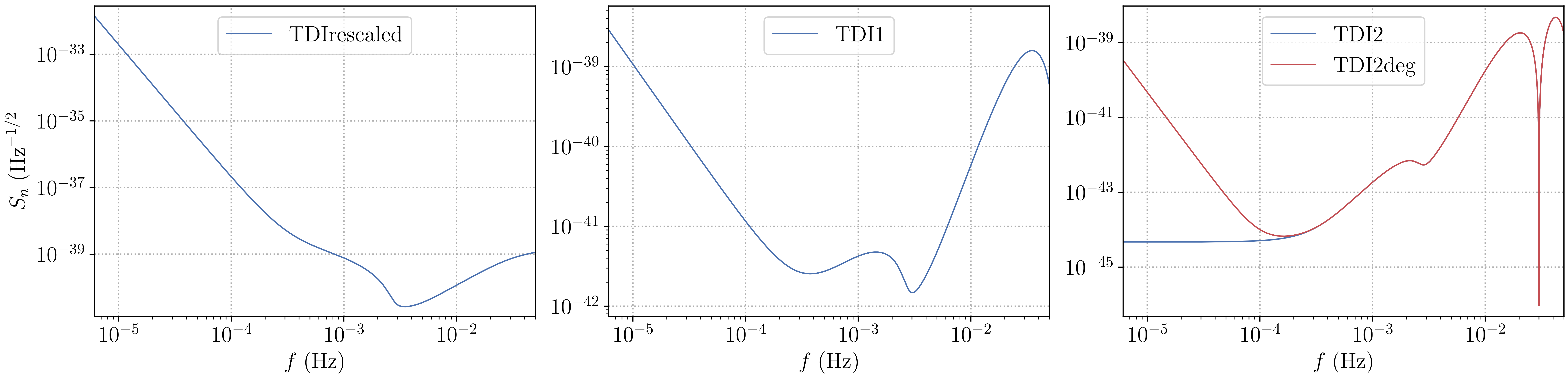
- $\Xi > 1$ represents how much information has been erased, notably by tapering (can also be <1 in some cases)
- Cramer-Rao constraint: $\Xi\Upsilon > 1$

Noise PSDs for LISA

TDIrescaled: $\tilde{y}_{slr} \sim \sin(\pi f L(1 - k \cdot n))\tilde{h}$

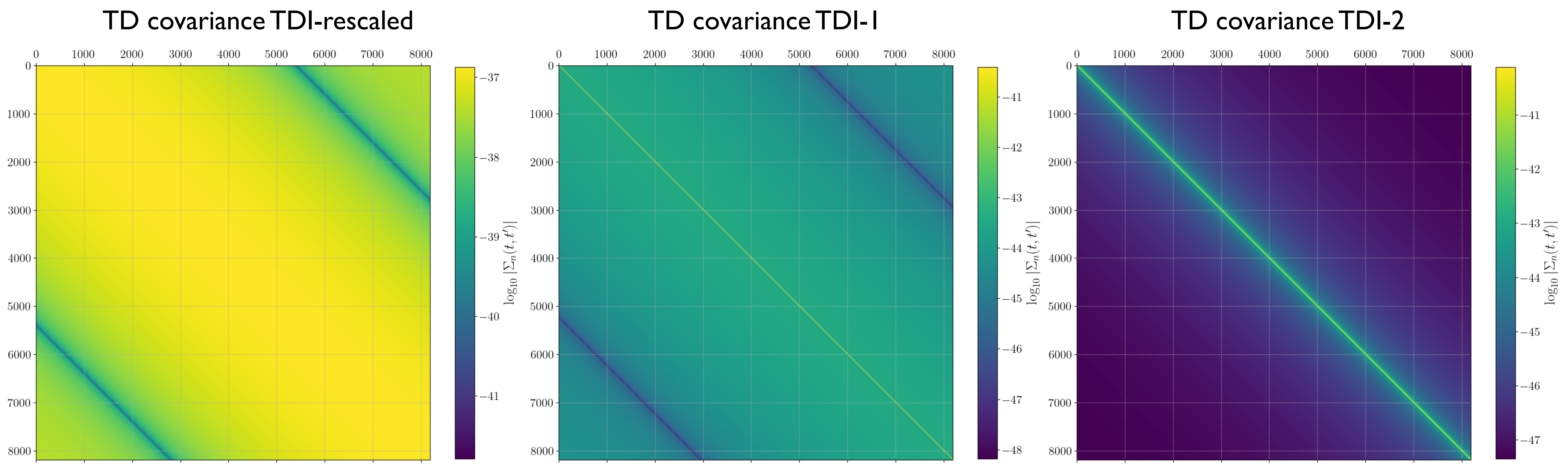
TDI1: $\tilde{X} \sim \sin(\pi f L(1 - k \cdot n))\sin(2\pi f L)\tilde{h}$

TDI2: $\tilde{X}_2 \sim \sin(\pi f L(1 - k \cdot n))\sin(2\pi f L)\sin(4\pi f L)\tilde{h}$

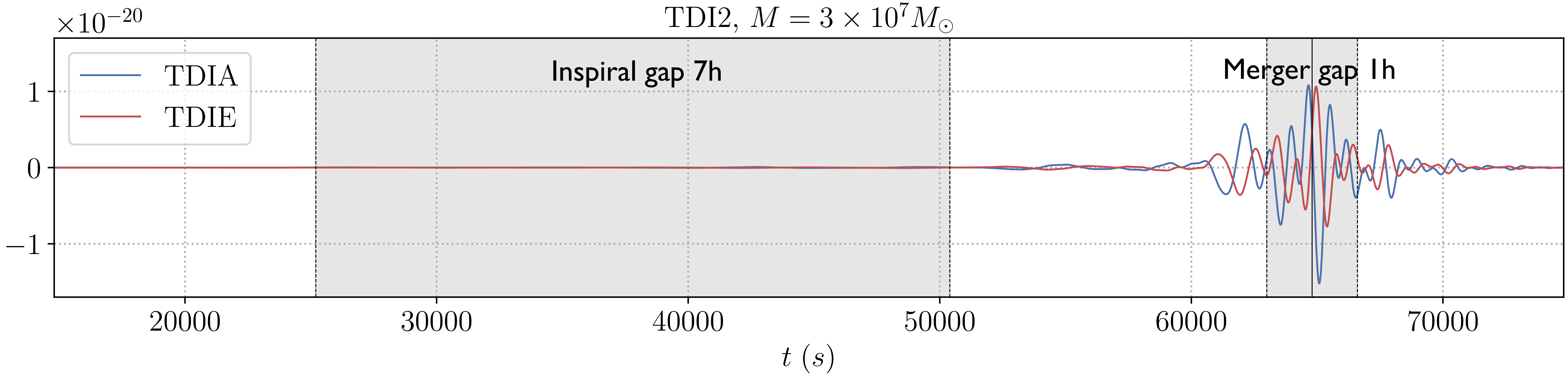


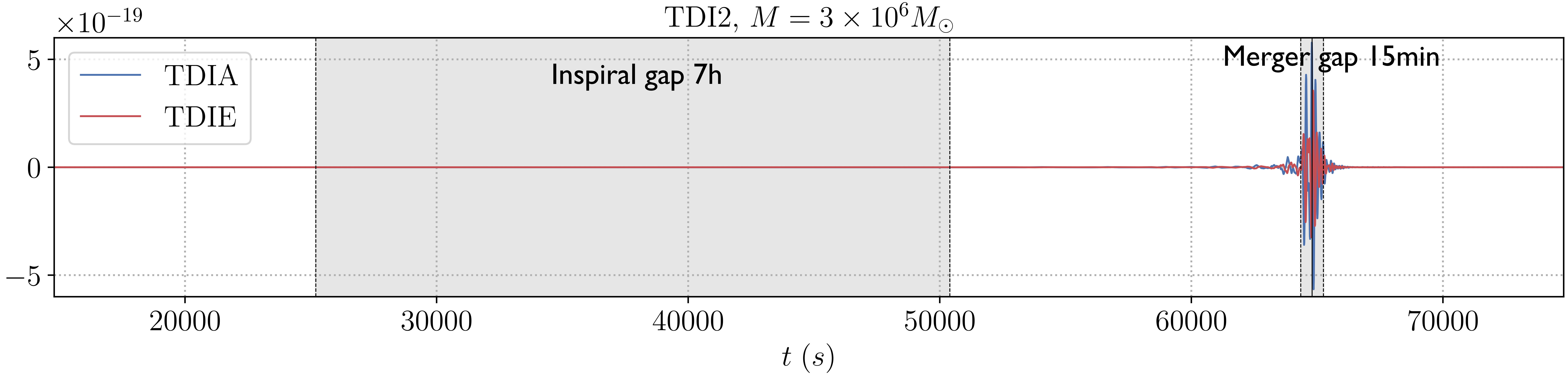
- Dependency on the sampling rate: a smaller Δt gives access to higher frequencies
- Dependency on the total duration: a smaller Δf gives access to lower frequencies

Covariance matrices in Time Domain



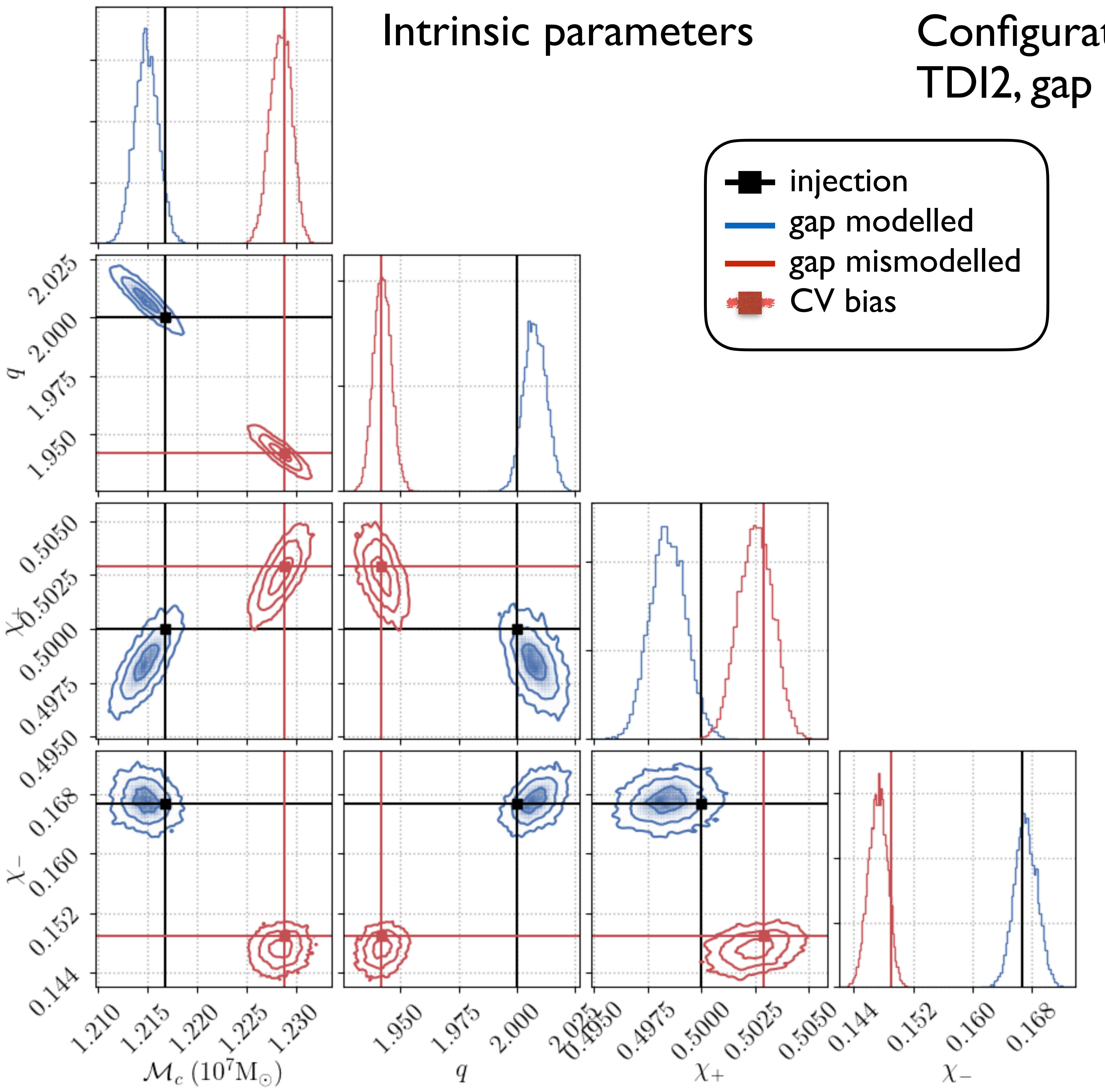
MBHBs: signals and gap configurations





Results: example PE, comparing with CV bias

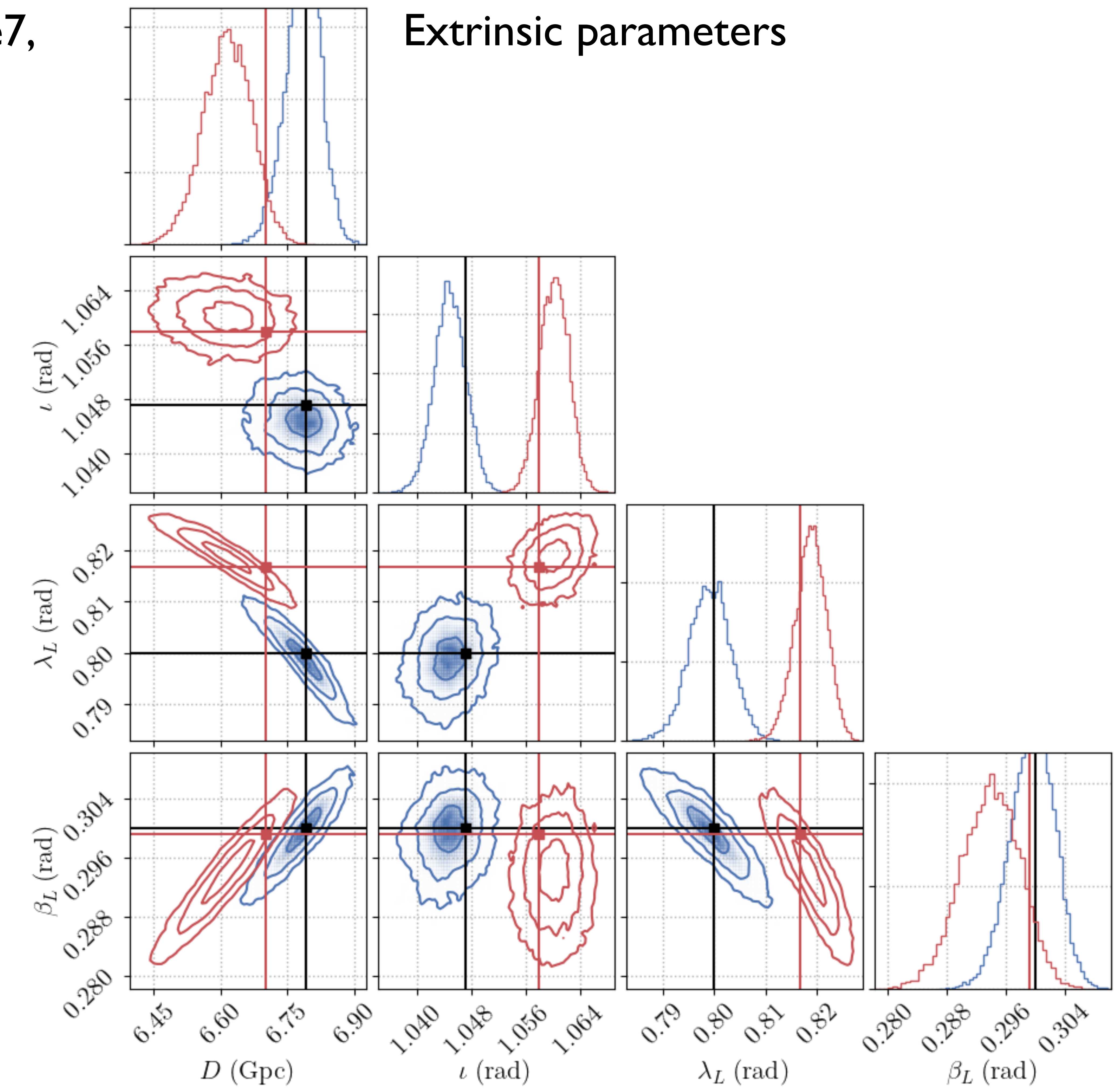
Intrinsic parameters



Configuration: M=3e7, TDI2, gap at merger

- injection
- gap modelled
- gap mismodelled
- CV bias

Extrinsic parameters

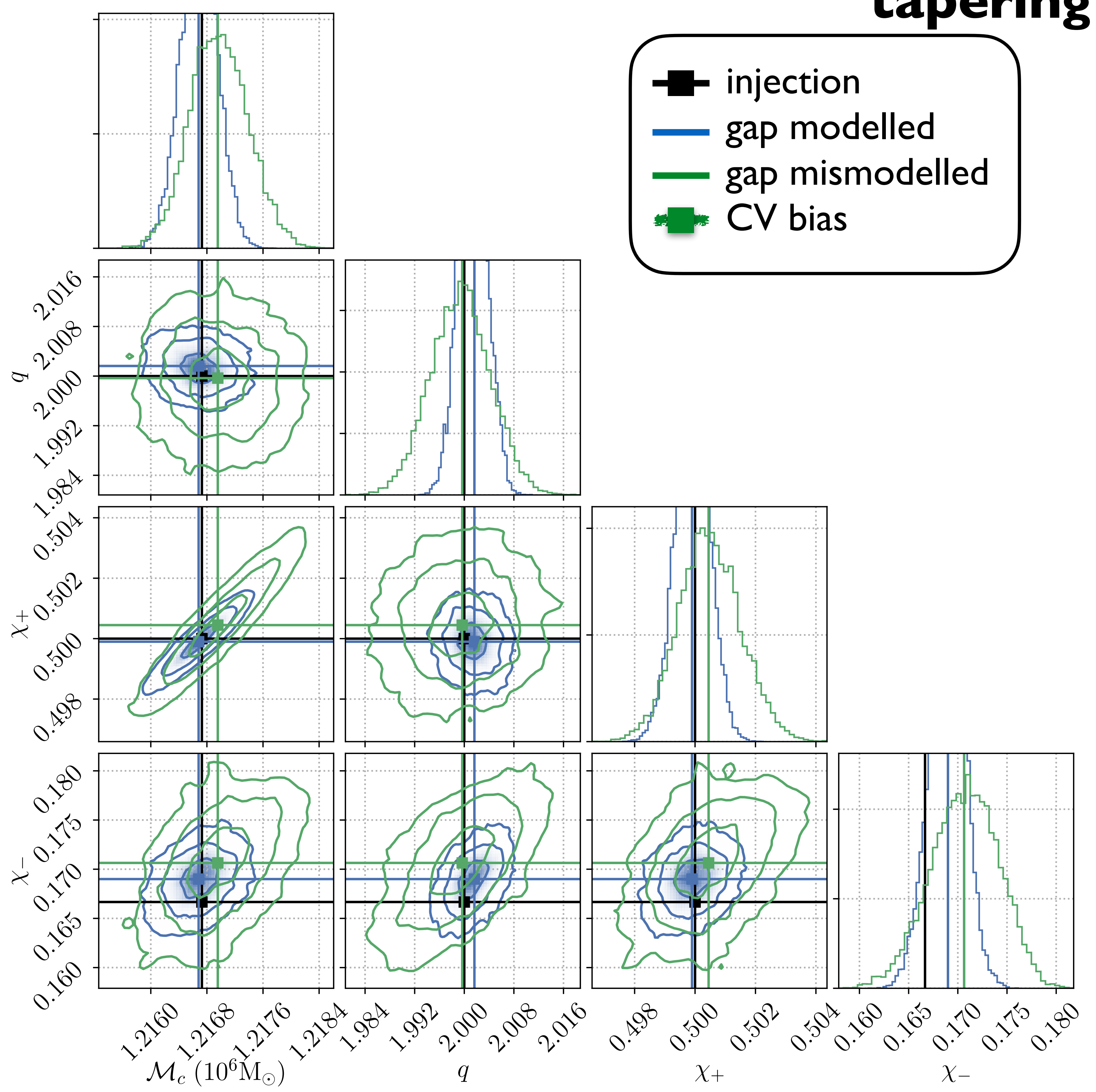


- Significant biases from mismodelling
- Bias well captured by CV

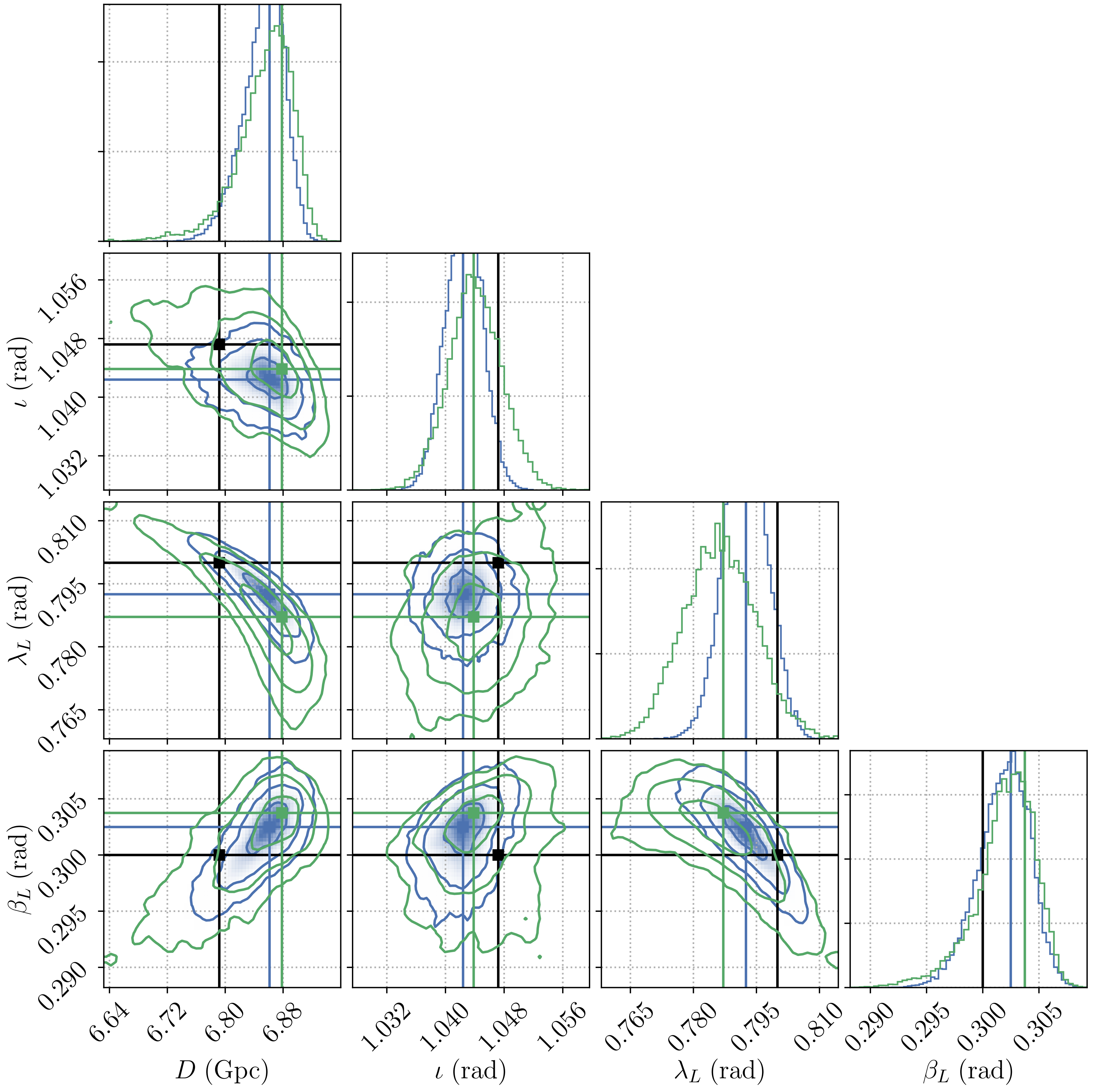
Results: example PE, comparing with CV bias

Configuration: M=3e6, TDI2, gap at merger, **tapering 10min**

Intrinsic parameters

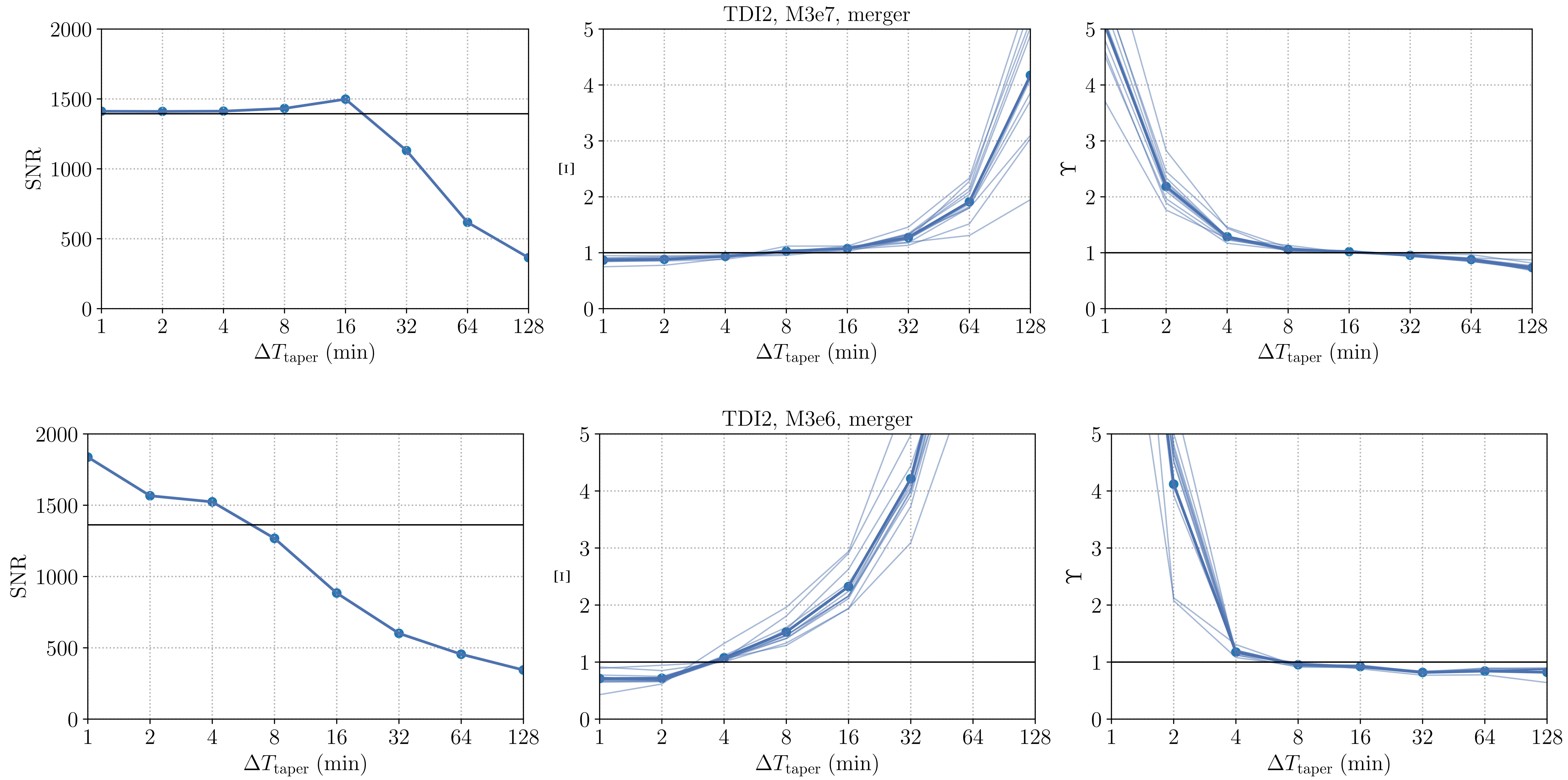


Extrinsic parameters



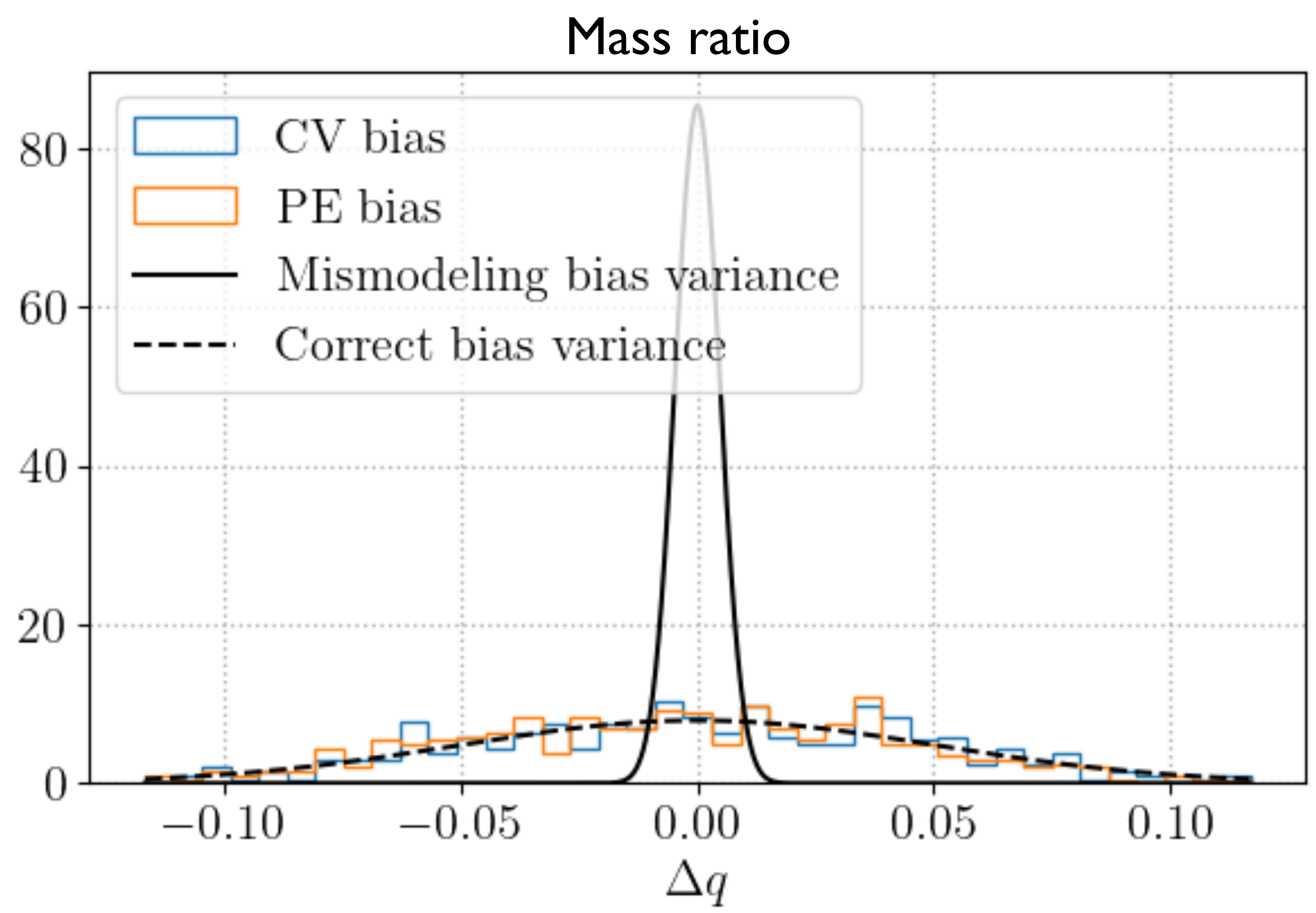
- Biases from mismodelling alleviated by tapering
- Loss of information

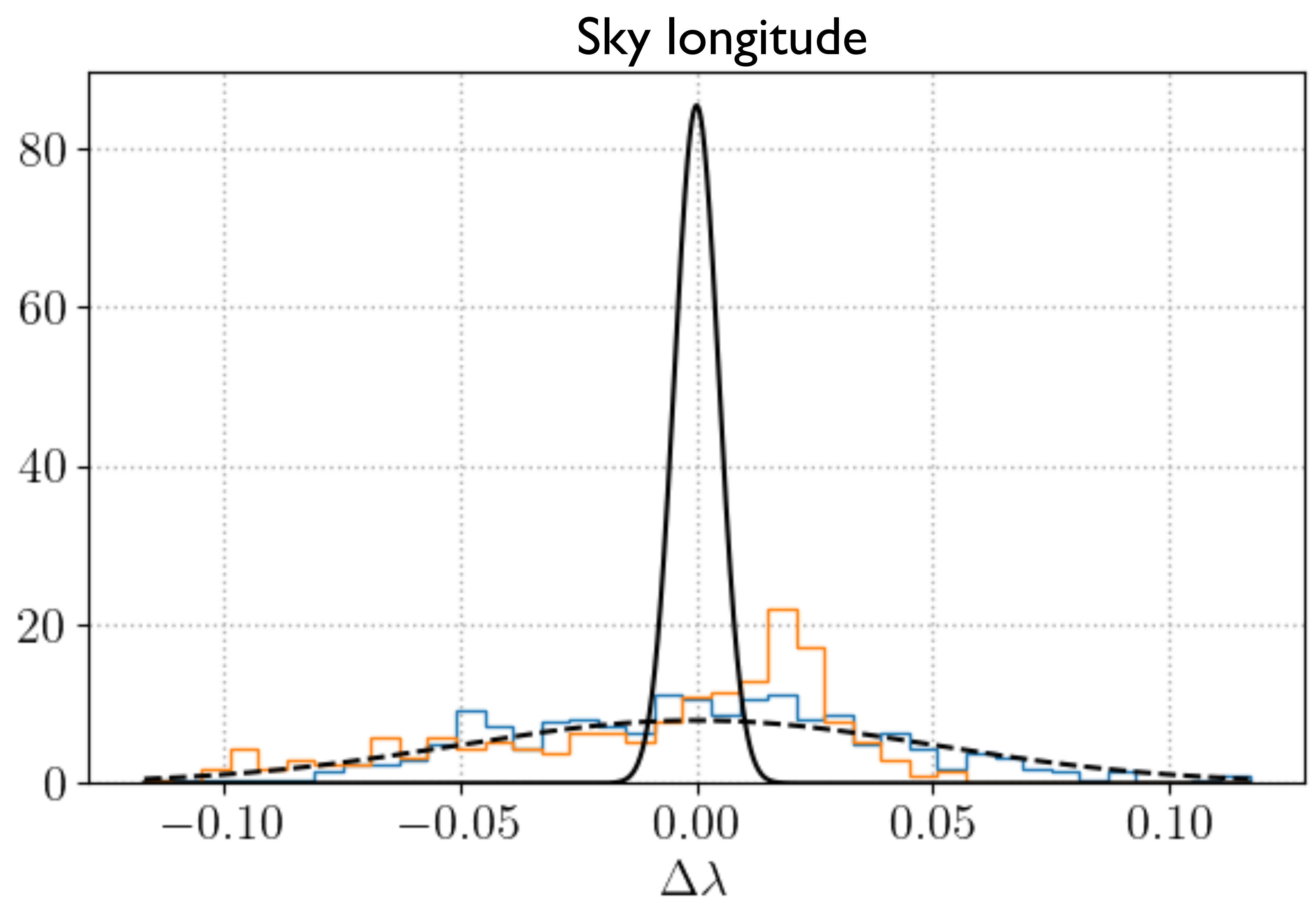
Influence of the tapering length



- Trade-off between statistical inconsistency and loss of information

Results: comparing CV mismodeling prediction to PE



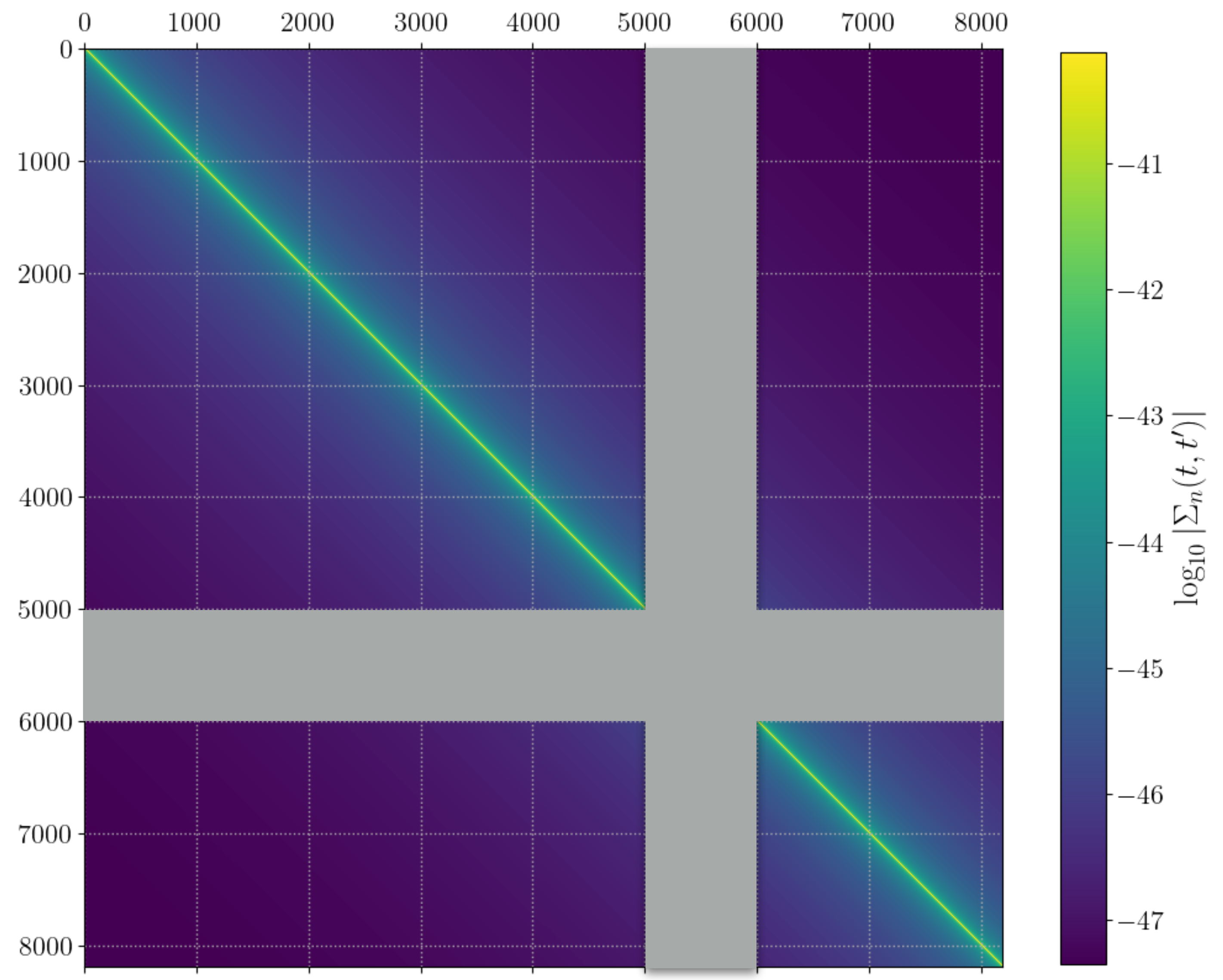


- Cutler-Vallisneri bias: calculated with the wrong covariance
- PE bias: from full PE runs, calculated with the wrong covariance
- Mismodelling bias variance: prediction $\langle \Delta\theta^i_{\mathrm{bf}'} \Delta\theta^j_{\mathrm{bf}'} \rangle$
- Correct bias variance: Fisher prediction with the correct covariance

Validation of the CV-inspired Υ to assess the impact of mismodelling noise

Results: biases from using the Whittle likelihood on data with gaps

Covariance with gap



- Correct modelling: pseudo-inverse of gated covariance

$$\ln \mathcal{L}_{\rm right} = -\frac{1}{2} X^T (W \Sigma W)^+ X$$

- Incorrect modelling: Whittle for gated data and signal

$$\ln \mathcal{L}_{\rm wrong} = -\frac{1}{2} X^T W \Sigma^{-1} W X$$

Mismodeling $\overline{\Upsilon}$ data gaps

		TDI2				TDI1				TDI0			
		Merger		Insp.		Merger		Insp.		Merger		Insp.	
Data	Model	M3e7	M3e6	M3e7	M3e6	M3e7	M3e6	M3e7	M3e6	M3e7	M3e6	M3e7	M3e6
Coherent, gap	Whittle, gated	12.2	40.5	1.3	1.4	1.2	1.6	1.0	1.0	39.7	38.0	21.9	6.4
Coherent, gap	Whittle, taper 10min	1.0	0.9	1.0	1.0	1.0	0.9	1.0	1.0	8.6	3.2	5.5	1.6
Coherent, gap	Whittle, taper 30min	1.0	0.8	1.0	1.0	1.0	0.8	1.0	1.0	6.1	2.0	3.0	1.1
Coherent, gap	Seg. Whittle, taper 10min	1.1	0.9	1.0	1.0	1.0	0.9	1.0	1.0	10.7	2.9	6.5	1.9
Coherent, gap	Seg. Whittle, taper 30min	1.0	0.8	1.0	1.0	1.0	0.8	1.0	1.0	6.6	2.0	3.6	1.4

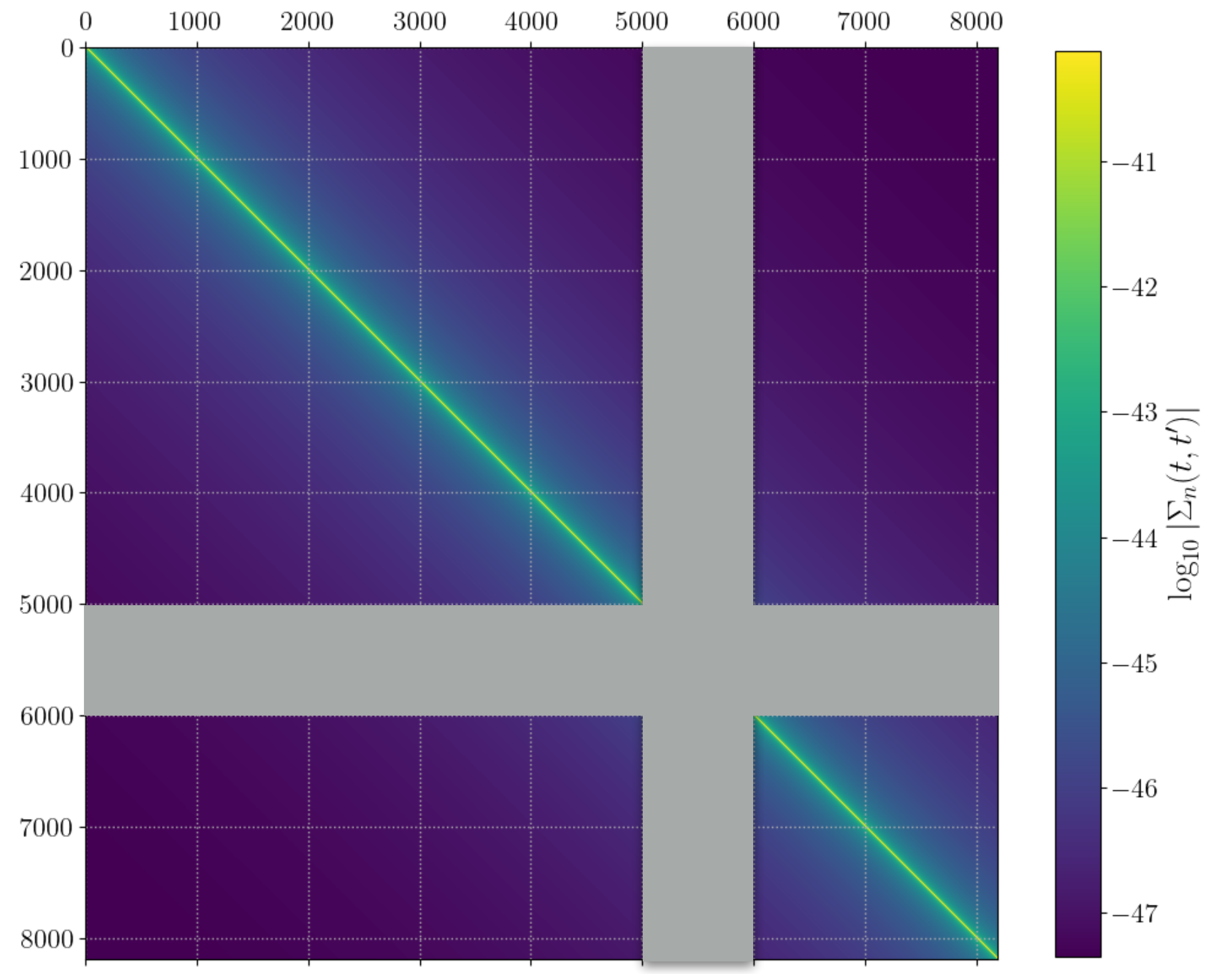
Mismodeling Ξ data gaps

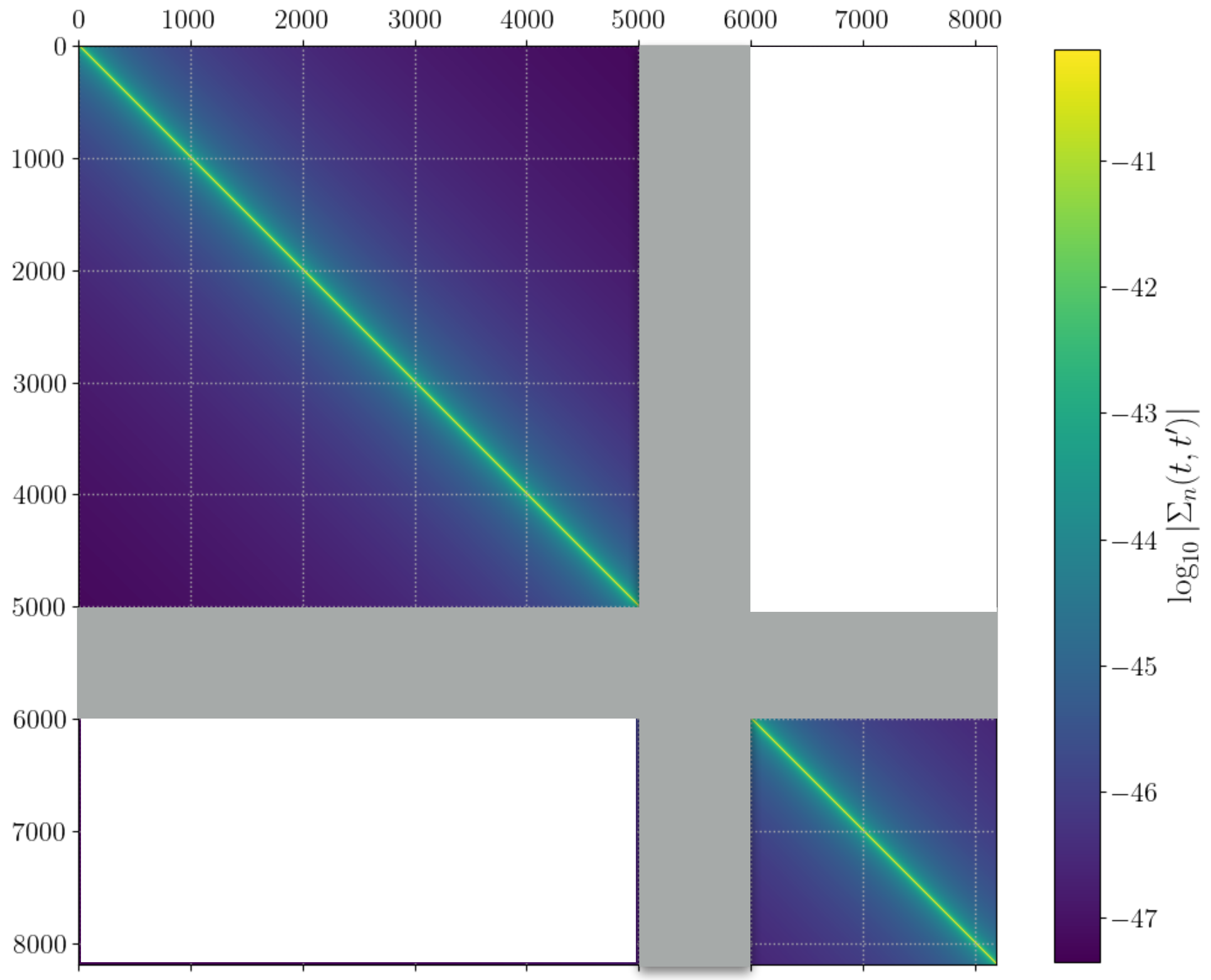
		TDI2				TDI1				TDI0			
		Merger		Insp.		Merger		Insp.		Merger		Insp.	
Data	Model	M3e7	M3e6	M3e7	M3e6	M3e7	M3e6	M3e7	M3e6	M3e7	M3e6	M3e7	M3e6
Coherent, gap	Whittle, gated	0.8	0.5	1.0	1.0	1.0	0.9	1.0	1.0	0.9	0.9	1.0	1.0
Coherent, gap	Whittle, taper 10min	1.0	1.7	1.0	1.0	1.1	1.8	1.0	1.0	1.1	1.7	1.0	1.0
Coherent, gap	Whittle, taper 30min	1.2	3.9	1.0	1.0	1.3	3.9	1.0	1.0	1.3	3.7	1.0	1.0
Coherent, gap	Seg. Whittle, taper 10min	1.0	1.8	1.0	1.0	1.1	1.8	1.0	1.0	1.1	1.8	1.0	1.0
Coherent, gap	Seg. Whittle, taper 30min	1.2	4.0	1.0	1.0	1.3	3.9	1.0	1.0	1.3	3.8	1.0	1.0

Question: how wrong is it to mismodel the covariance using Whittle ? Effect of tapering ?

Results: assumptions about segment independence

Covariance with gap, correlated vs uncorrelated segments



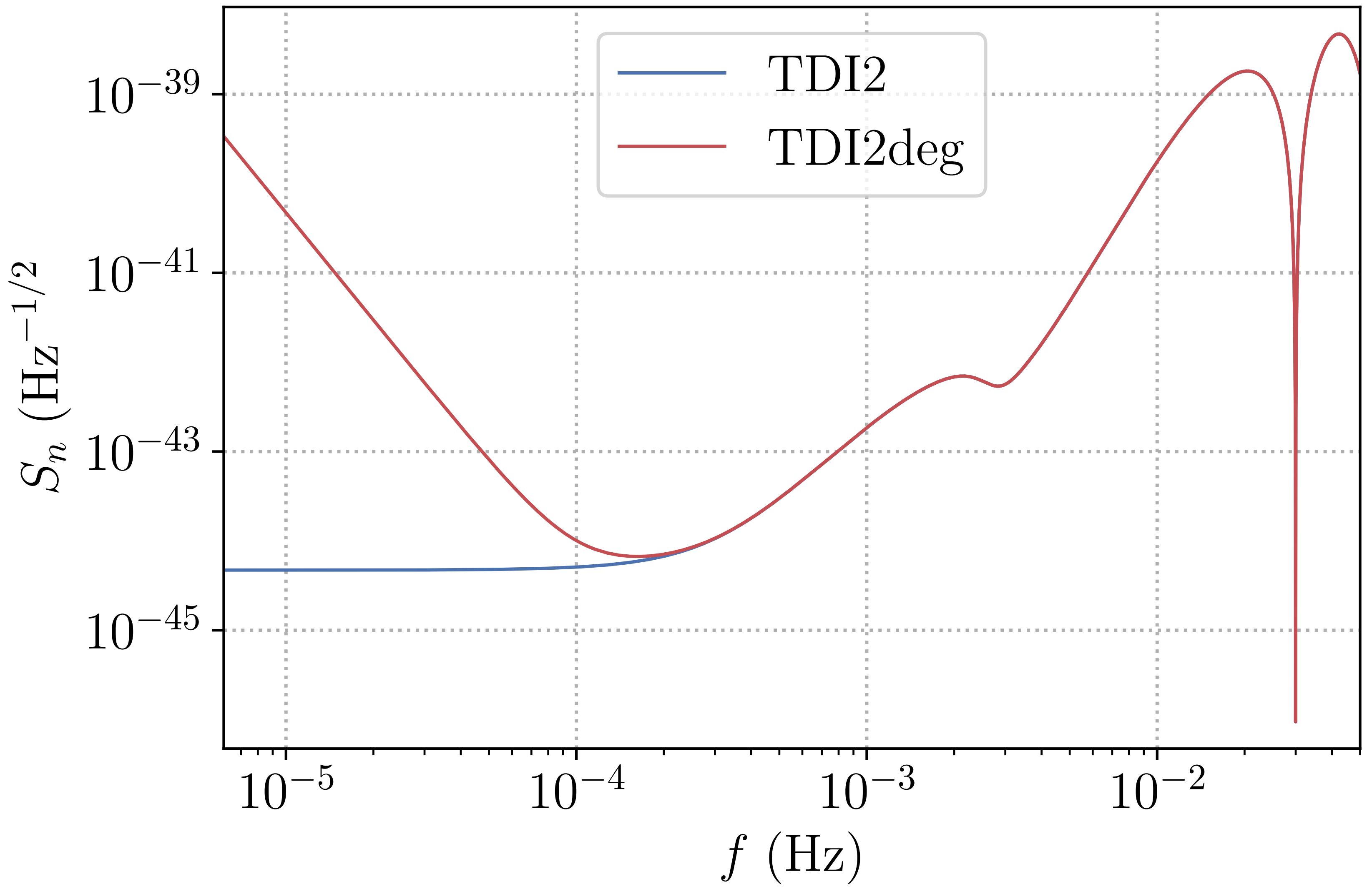


Question: how wrong is it to mismodel data segments as correlated/independent ?

Mismodeling $\overline{\Upsilon}$ segment independence		TDI2				TDI1				TDI0			
		Merger		Insp.		Merger		Insp.		Merger		Insp.	
Data	Model	M3e7	M3e6	M3e7	M3e6	M3e7	M3e6	M3e7	M3e6	M3e7	M3e6	M3e7	M3e6
Coherent, gap	Incoherent, gap	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Incoherent, gap	Coherent, gap	1.2	1.1	1.1	1.5	1.0	1.0	1.0	1.0	26.2	8.1	13.4	14.2
Coherent, split	Incoherent, split	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Incoherent, split	Coherent, split	5.6	2.1	1.3	1.5	1.3	1.6	1.0	1.0	27.3	41.0	11.8	11.9

Mismodeling $\overline{\Xi}$ segment independence		TDI2				TDI1				TDI0			
		Merger		Insp.		Merger		Insp.		Merger		Insp.	
Data	Model	M3e7	M3e6	M3e7	M3e6	M3e7	M3e6	M3e7	M3e6	M3e7	M3e6	M3e7	M3e6
Coherent, gap	Incoherent, gap	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Incoherent, gap	Coherent, gap	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Coherent, split	Incoherent, split	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.1	1.0	1.0
Incoherent, split	Coherent, split	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.9	1.0	1.0

Results: mismodeling of low-frequency noise



For a stationary process (no gap), the effect should be weak by frequency independence.

For a non-stationary process (gap), can the mismodeling error at low frequencies affect higher frequencies where the signal's SNR is ?

Question: how important is it to model the PSD at low frequencies (which is challenging) ?

Mismodeling $\overline{\Upsilon}$ low-f PSD

Data	Model	TDI2 Merger M3e7	TDI2 Merger M3e6	TDI2 Insp. M3e7	TDI2 Insp. M3e6
PSD TDI2$_{\rm red}$, no gap	PSD TDI2, no gap	1.0	1.0	1.0	1.0
PSD TDI2$_{\rm red}$, gap	PSD TDI2, gap	2.2	1.3	1.2	1.1
PSD TDI2$_{\rm red}$, gap	PSD TDI2, Whittle, gated	13.4	40.9	1.6	1.4
PSD TDI2$_{\rm red}$, gap	PSD TDI2, Whittle, taper 10min	3.6	1.2	1.3	1.1
PSD TDI2$_{\rm red}$, gap	PSD TDI2, Whittle, taper 30min	2.2	1.0	1.2	1.1
PSD TDI2$_{\rm red}$, gap	PSD TDI2$_{\rm red}$, Whittle, gated	9.6	39.0	1.2	1.2
PSD TDI2$_{\rm red}$, gap	PSD TDI2$_{\rm red}$, Whittle, taper 10min	1.6	1.3	1.0	1.0
PSD TDI2$_{\rm red}$, gap	PSD TDI2$_{\rm red}$, Whittle, taper 30min	1.2	1.1	1.0	1.0

Mismodeling $\overline{\Xi}$ low-f PSD

Data	Model	TDI2 Merger M3e7	TDI2 Merger M3e6	TDI2 Insp. M3e7	TDI2 Insp. M3e6
PSD TDI2$_{\rm red}$, no gap	PSD TDI2, no gap	1.0	1.0	1.0	1.0
PSD TDI2$_{\rm red}$, gap	PSD TDI2, gap	1.0	1.0	1.0	1.0
PSD TDI2$_{\rm red}$, gap	PSD TDI2, Whittle, gated	0.8	0.5	1.0	1.0
PSD TDI2$_{\rm red}$, gap	PSD TDI2, Whittle, taper 10min	1.0	1.7	1.0	1.0
PSD TDI2$_{\rm red}$, gap	PSD TDI2, Whittle, taper 30min	1.2	3.9	1.0	1.0
PSD TDI2$_{\rm red}$, gap	PSD TDI2$_{\rm red}$, Whittle, gated	0.9	0.5	1.0	1.0
PSD TDI2$_{\rm red}$, gap	PSD TDI2$_{\rm red}$, Whittle, taper 10min	1.1	1.7	1.0	1.0
PSD TDI2$_{\rm red}$, gap	PSD TDI2$_{\rm red}$, Whittle, taper 30min	1.2	3.9	1.0	1.0

Results and outlook

Results

- Derived a framework handling missing data in time or frequency domain — **caveat**: for short segments only
- Derived a measure of inconsistency for the scatter of best-fit parameters
- Application to the LISA case: exploration of different mismodelling settings
- This framework is a test-bed: allows to test assumptions, possibility to compare to e.g. imputation

Outlook

- Longer signals ? Low-mass MBHBs, EMRIs ?
- Extension to other sources of non-stationarity
- More exploration to be done…

The basics: Stationary Gaussian process

- Assumption: noise as **Gaussian process** described by its covariance

$$C(t,t') = \langle n(t)n(t') \rangle$$

- Assumption: underlying noise (before introducing gaps) **Stationary**, with autocorrelation depending on lag only

$$C(t,t') = C(0,t'-t) \equiv C(t'-t)$$

- In the Fourier domain, this leads to independence:

$$\langle \tilde{n}(f)\tilde{n}^*(f') \rangle = \frac{1}{2} S_n(f) \delta(f-f')$$

- Usually represented with 1-sided PSD $S_n(f)$

- Noise-weighted inner product over positive frequencies:

$$(a|b) = 4\mathrm{Re} \int \frac{df}{S_n(f)} \tilde{a}^*(f)\tilde{b}(f)$$

- Likelihood:

$$\ln \mathcal{L}(\theta) = -\frac{1}{2}(h(\theta)-d|h(\theta)-d)$$

Whittle likelihood

The Fourier-domain covariance matrix is diagonal: from $N \times N$ to N !

TD/FD data vectors and covariance

Time domain covariance matrix

$$\langle N N^T \rangle = \Sigma$$

Sationarity imposes a Toeplitz structure

$$C(t, t') \equiv C(t - t')$$

$$A = \begin{bmatrix} a_0 & a_{-1} & a_{-2} & \cdots & \cdots & a_{-(n-1)} \\ a_1 & a_0 & a_{-1} & \ddots & & \vdots \\ a_2 & a_1 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & a_{-1} & a_{-2} \\ \vdots & & \ddots & a_1 & a_0 & a_{-1} \\ a_{n-1} & \cdots & \cdots & a_2 & a_1 & a_0 \end{bmatrix}$$

Diagonality after DFT requires in fact **Circulant** structure (periodicity)

$$C = \begin{bmatrix} c_0 & c_{n-1} & \cdots & c_2 & c_1 \\ c_1 & c_0 & c_{n-1} & & c_2 \\ \vdots & c_1 & c_0 & \ddots & \vdots \\ c_{n-2} & & \ddots & \ddots & c_{n-1} \\ c_{n-1} & c_{n-2} & \cdots & c_1 & c_0 \end{bmatrix}$$

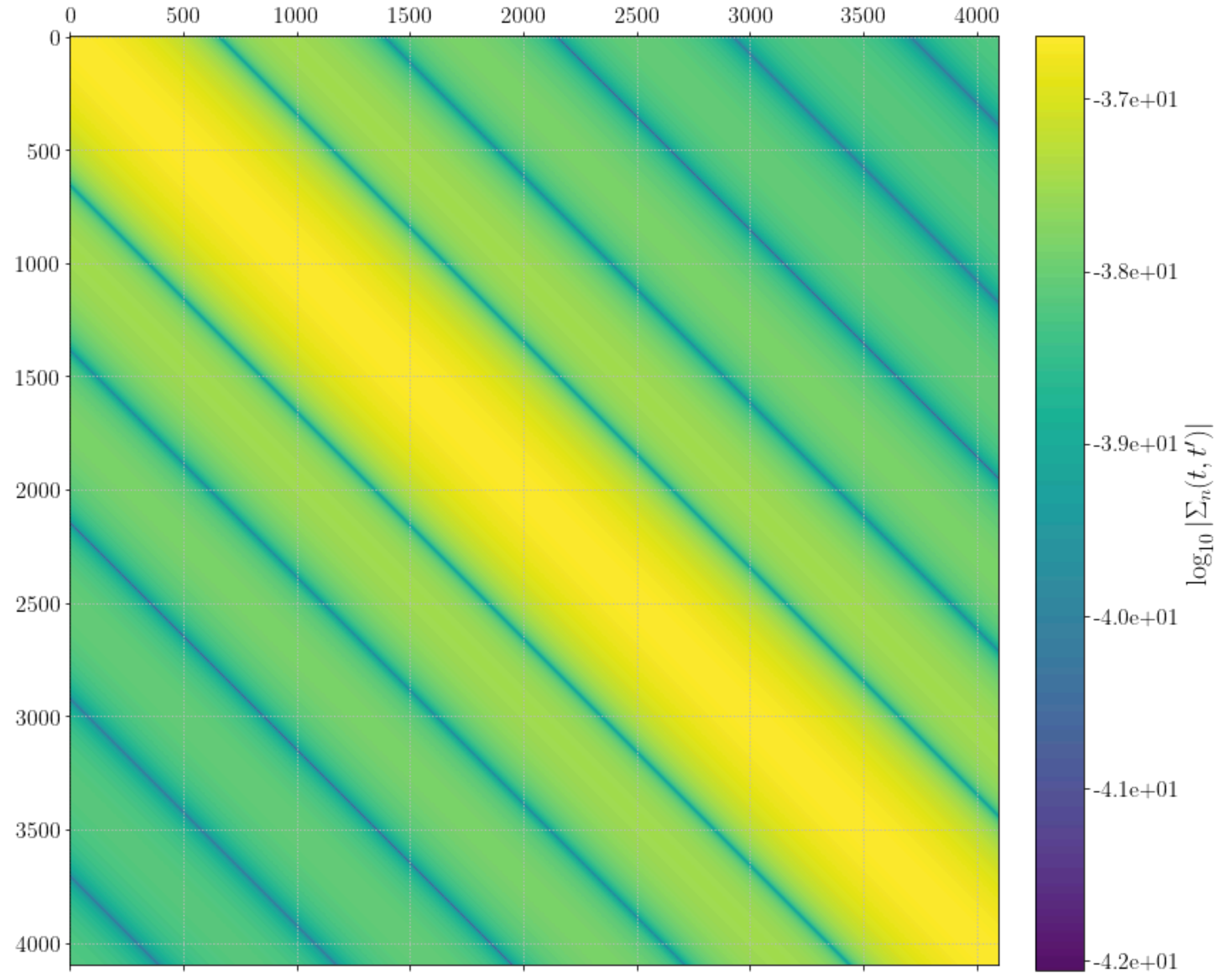
Discrete Fourier transform

$$\tilde{F}(f) = \int dt\, e^{-2i\pi f t} F(t)$$

$$\tilde{F}(f_j) = \Delta t \sum_{i=0}^{N-1} \omega^{-ij} F(t_i) \qquad \omega = e^{\frac{2i\pi}{N}}$$

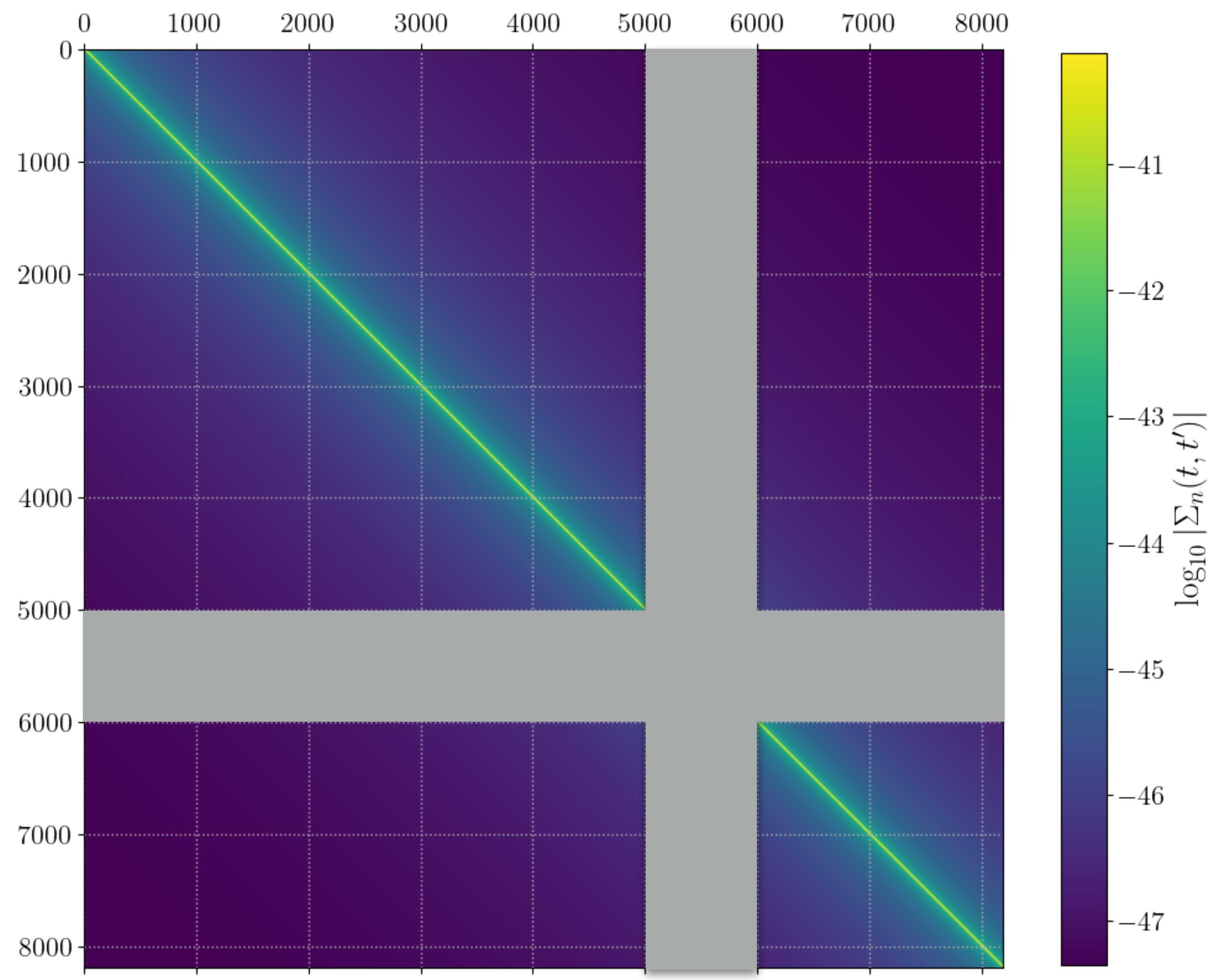
$$\langle \tilde{N} \tilde{N}^\dagger \rangle = \tilde{\Sigma}$$ with $\tilde{\Sigma}$ diagonal

Example TD covariance



Marginalizing out the missing data: Time Domain

Direct marginalization:



Truncation of the Gaussian covariance:

$$X = H(\theta) - D = -N \qquad X = (X_1, X_2)$$

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12}^T & \Sigma_{22} \end{pmatrix}$$

$$\mathcal{L}_{\mathrm{gap}} = p(N_1) = \int dN_2\, p(N_1, N_2)$$

$$\ln \mathcal{L}_{\mathrm{gap}} = -\frac{1}{2} X_1^T \Sigma_{11}^{-1} X_1$$

Gap seen as a gated process:

For an invertible modulation W:

$$X \to WX$$

$$X^T \Sigma^{-1} X = (WX)^T (W\Sigma W)^{-1} (WX)$$

For an gating function W, non-invertible: $W = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & 0 \end{pmatrix}$

Moore-Penrose **pseudo-inverse** (uniquely defined):

$$A^+ A A^+ = A^+$$

$$A A^+ A = A$$

$$(AA^+)^\dagger = AA^+$$

$$(A^+A)^\dagger = A^+A$$

$$(W\Sigma W)^+ = \begin{pmatrix} \Sigma_{11}^{-1} & 0 \\ 0 & 0 \end{pmatrix}$$

Equivalence with the marginalized likelihood:

$$X^T (W\Sigma W)^+ X = X_1^T \Sigma_{11}^{-1} X_1$$

Note that windowing the data becomes facultative !

$$(WX)^T (W\Sigma W)^+ (WX) = X^T (W\Sigma W)^+ X$$

Marginalizing out the missing data: Fourier Domain

DFT in linear algebra:

$$\tilde{F}(f) = \int dt\, e^{-2i\pi f t} F(t)$$

$$\tilde{F}(f_j) = \Delta t \sum_{i=0}^{N-1} \omega^{-ij} F(t_i) \quad \omega = e^{\frac{2i\pi}{N}}$$

Using the DFT matrix P:

$$\tilde{F} = \Delta t \sqrt{N} P F$$

$$P_{ij} = \frac{1}{\sqrt{N}} \omega^{-ij}$$

$$P P^{\dagger} = \mathbb{1}$$

FD covariance and FD window:

$$\tilde{\Sigma} = N \Delta t^2 P \Sigma P^{\dagger}$$

$$\tilde{W} = P W P^{\dagger}$$

Translation from time to Fourier domain:

The pseudo-inverse is transparent to the unitary matrix P:

$$P \left(W \Sigma W\right)^{+} P^{\dagger} = \left(P W \Sigma W P^{\dagger}\right)^{+}$$

This allows for a direct translation:

$$\ln \mathcal{L}_{\mathrm{gap}} = -\frac{1}{2} (WX)^T \left(W \Sigma W\right)^{+} (WX) = -\frac{1}{2} \left(\widetilde{WX}\right)^{\dagger} \left(\tilde{W} \tilde{\Sigma} \tilde{W}\right)^{+} \left(\widetilde{WX}\right)$$

Windowing the data is facultative, in FD also:

$$\ln \mathcal{L}_{\mathrm{gap}} = -\frac{1}{2} \left(\widetilde{WD} - \tilde{H}\right)^{\dagger} \left(\tilde{W} \tilde{\Sigma} \tilde{W}\right)^{+} \left(\widetilde{WD} - \tilde{H}\right)$$

In practice, computing the FD gated covariance matrix is a convolution:

$$\left(\tilde{W} \tilde{\Sigma} \tilde{W}\right)_{ij} = \frac{\Delta f}{2} \sum_k S_n^k \tilde{w}_{i-k} \tilde{w}^*_{j-k}$$

cost scales as $\mathcal{O}(N^2 \log N)$

$$\tilde{\Sigma}_w(f, f') = \frac{1}{2} \int_{-\infty}^{+\infty} dv\, S_n(v) \tilde{w}(f - v) \tilde{w}^*(f' - v)$$

Numerics: pseudo-inverse of the FD covariance

- The Moore-Penrose pseudo-inverse is uniquely defined
- Most straightfoward algorithm: use the **singular value decomposition** (SVD)

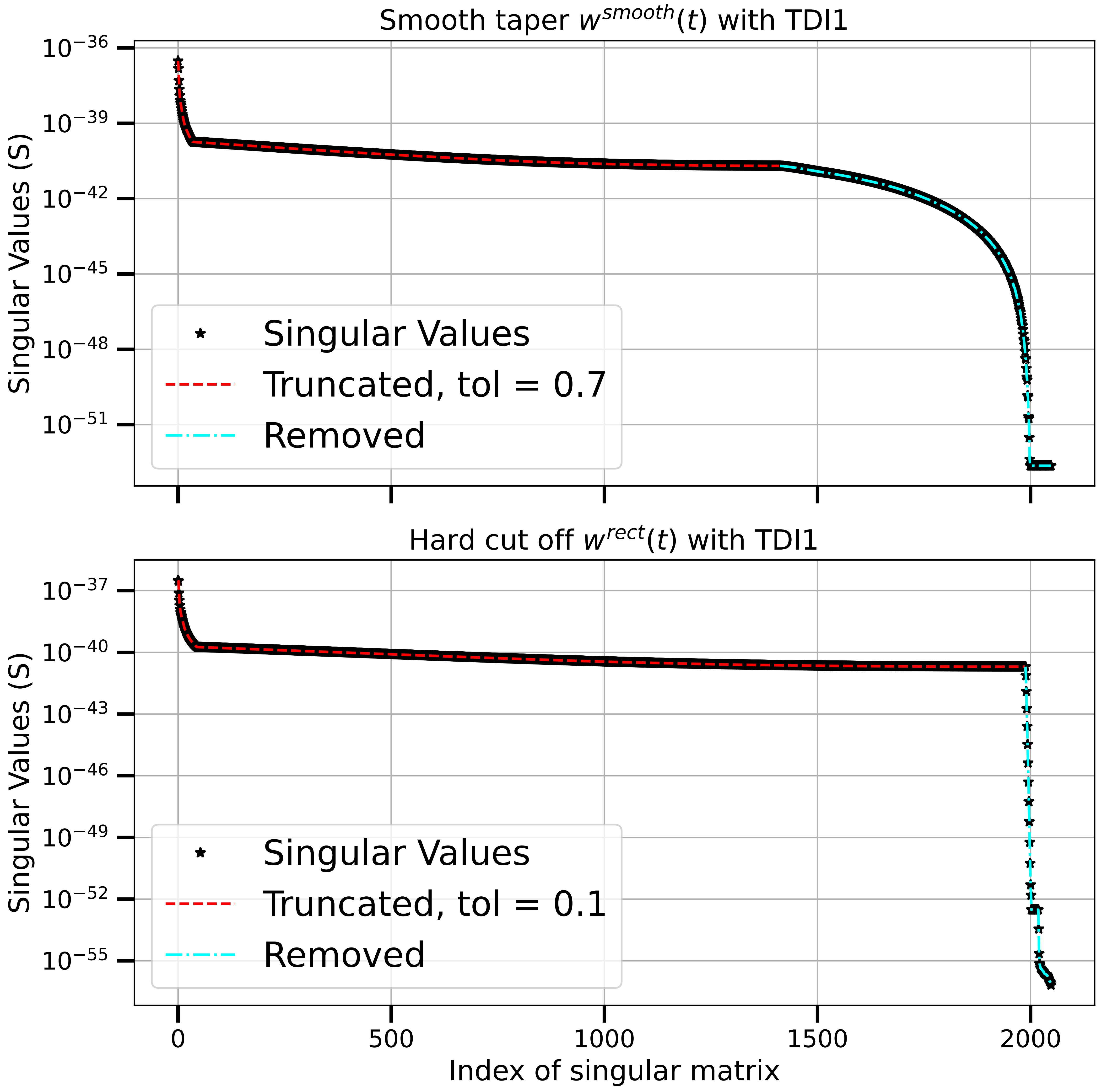
$$M = U\Sigma V^\dagger$$

$$\Sigma = (\sigma_1^2, \ldots, \sigma_p^2, 0, \ldots, 0)$$

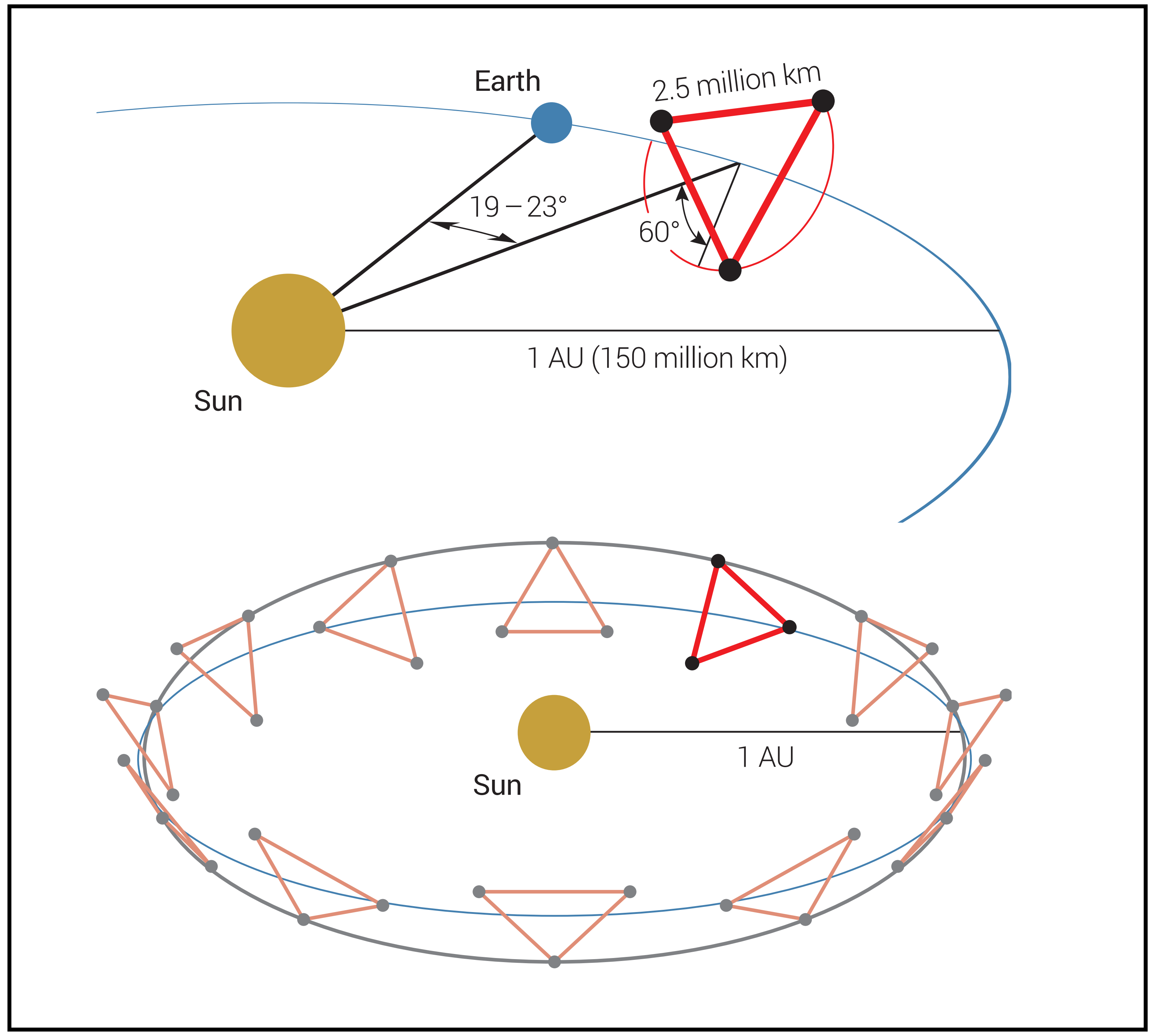
And use the pseudo-inverse of singular values:

$$\Sigma^+ = (1/\sigma_1^2, \ldots, 1/\sigma_p^2, 0, \ldots, 0)$$

- The procedure also allows for regularizing the covariance (for instance when dealing with smooth windows)



LISA data and TDI generations



Doppler delay from orbit, change in orientation

Analogous to 2 LIGO in motion at low frequencies only

One-link observables: from spacecraft s to spacecraft r through link s:

$$y = \Delta\nu/\nu$$

$$y_{slr} = \frac{1}{2}\frac{1}{1-\hat{k}\cdot n_l} n_l \cdot (h(t_s) - h(t_r)) \cdot n_l$$

Response time and frequency-dependent:

$$\mathcal{T}_{slr} = \frac{i\pi f L}{2} \mathrm{sinc}\left[\pi f L \left(1 - k\cdot n_l\right)\right] \exp\left[i\pi f \left(L + k\cdot(p_r + p_s)\right)\right] n_l \cdot P \cdot n_l(t_f)$$

\+ Time-delay interferometry (TDI), 1st and 2nd generations:

$$X_2^{GW} = \underbrace{\left[(y_{31}^{\mathrm{GW}} + y_{13,2}^{\mathrm{GW}}) + (y_{21}^{\mathrm{GW}} + y_{12,3}^{\mathrm{GW}})_{,22} - (y_{21}^{\mathrm{GW}} + y_{12,3}^{\mathrm{GW}}) - (y_{31}^{\mathrm{GW}} + y_{13,2}^{\mathrm{GW}})_{,33}\right]}_{X^{\mathrm{GW}}(t)}$$
$$- \underbrace{\left[(y_{31}^{\mathrm{GW}} + y_{13,2}^{\mathrm{GW}}) + (y_{21}^{\mathrm{GW}} + y_{12,3}^{\mathrm{GW}})_{,22} - (y_{21}^{\mathrm{GW}} + y_{12,3}^{\mathrm{GW}}) - (y_{31}^{\mathrm{GW}} + y_{13,2}^{\mathrm{GW}})_{,33}\right]_{,2233}}_{X^{\mathrm{GW}}(t-2L_2-2L_3)\simeq X^{\mathrm{GW}}(t-4L)} .$$

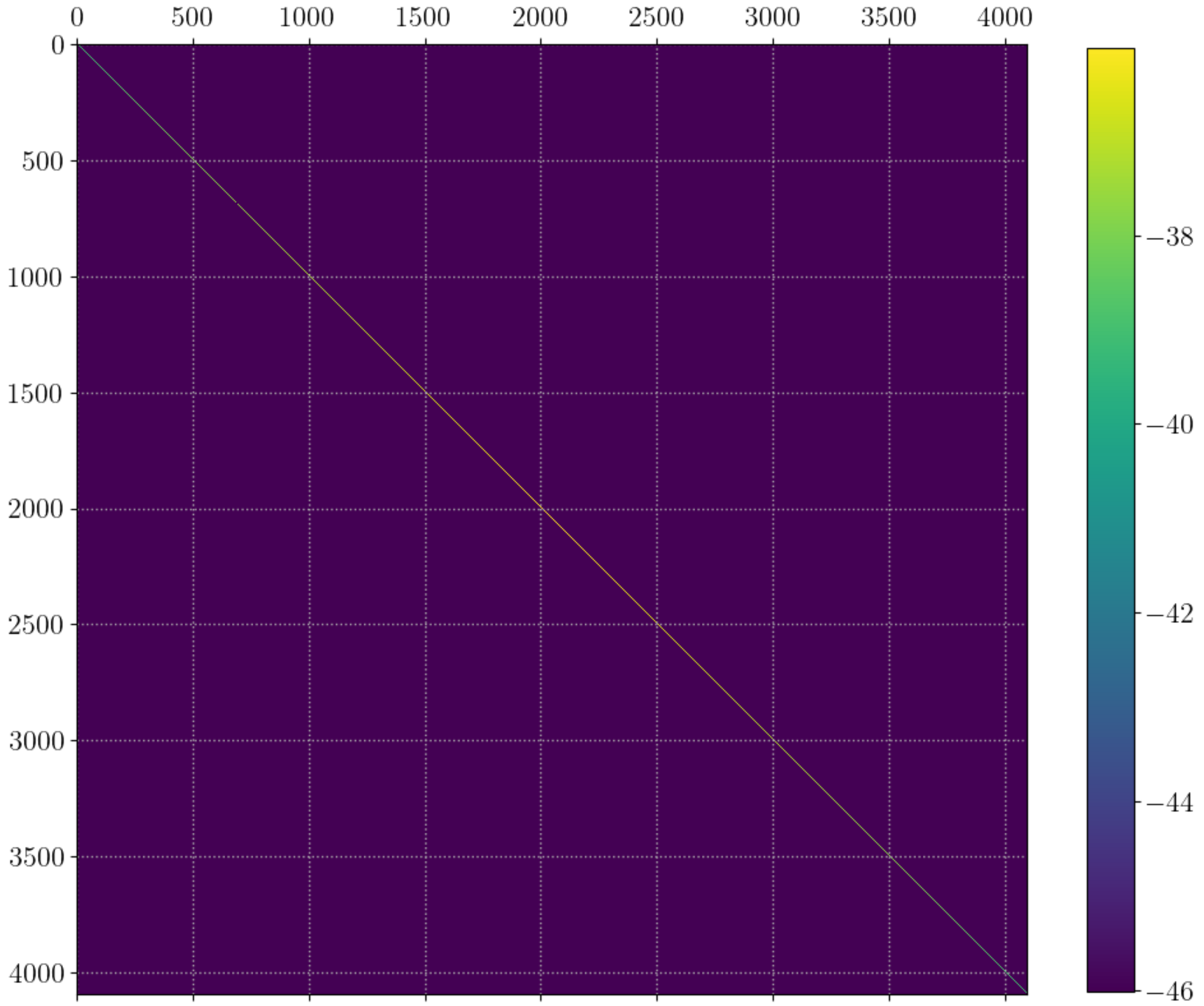
TDI variables amount to taking (discretized) time derivatives of the GW:

$$y_{slr} \sim h(t) - h(t-L)$$
$$X \sim y_{slr}(t) - y_{slr}(t-2L)$$
$$X_2 \sim X(t) - X(t-4L)$$

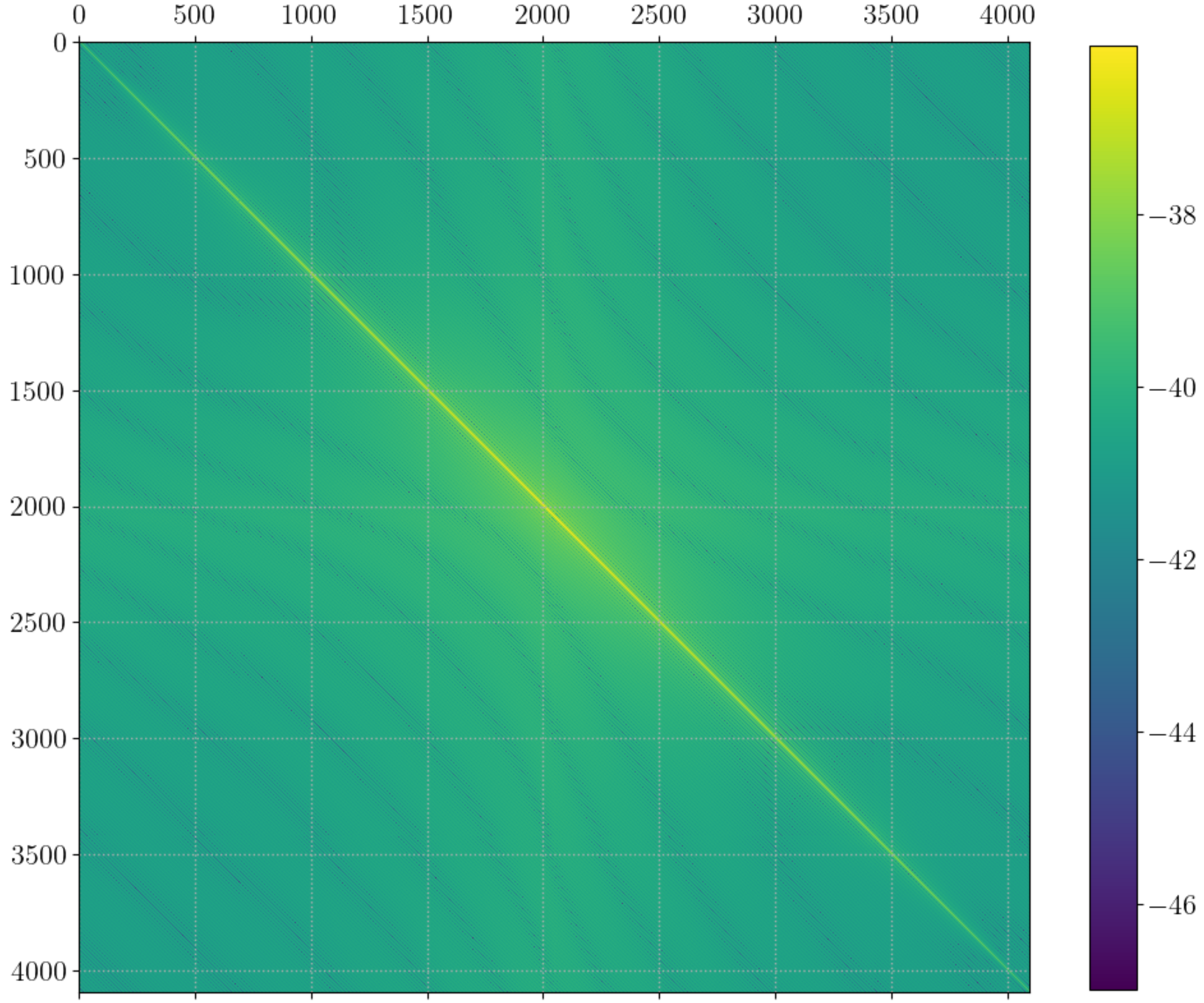
$$\tilde{y}_{slr} \sim \sin(\pi f L(1 - k\cdot n))\tilde{h}$$
$$\tilde{X} \sim \sin(\pi f L(1 - k\cdot n))\sin(2\pi f L)\tilde{h}$$
$$\tilde{X}_2 \sim \sin(\pi f L(1 - k\cdot n))\sin(2\pi f L)\sin(4\pi f L)\tilde{h}$$

Covariance matrices in Fourier Domain

FD covariance TDI-2 without gap (stationary)



FD covariance TDI-2 with gap



Note: need to consider both positive and negative frequencies

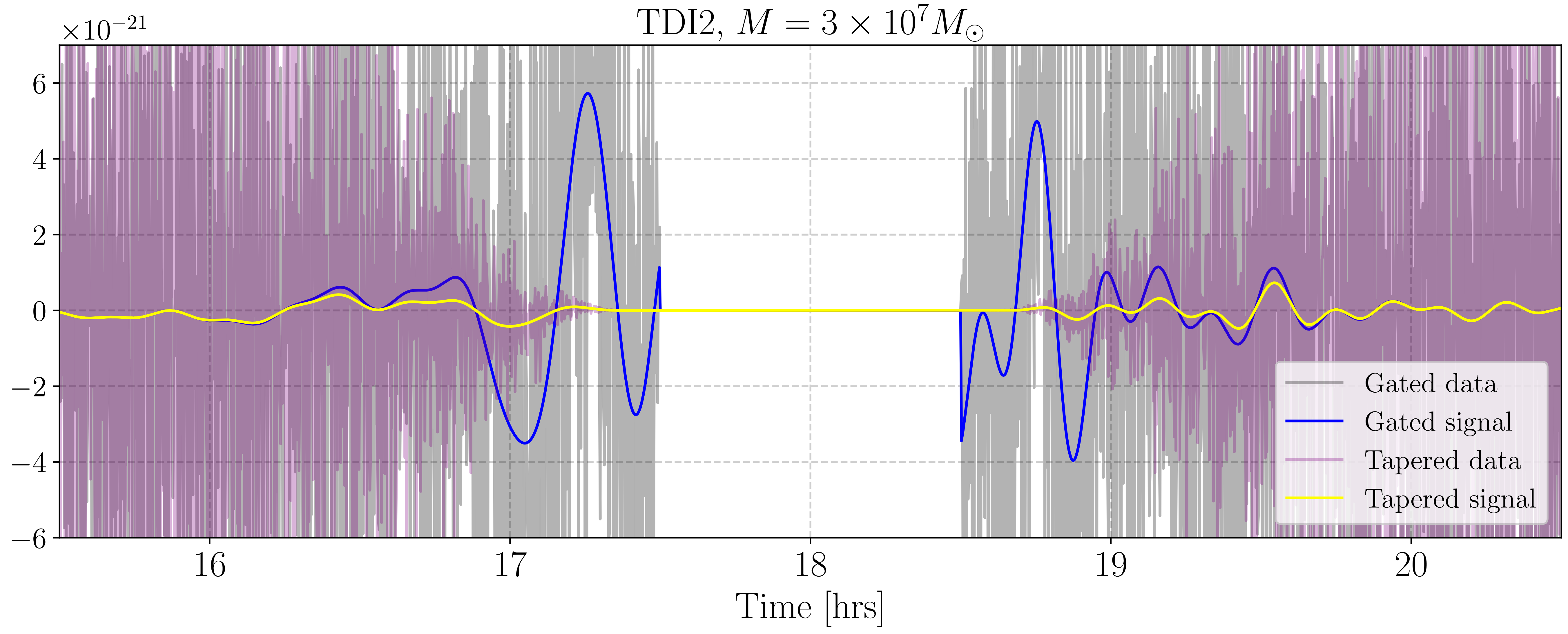
Windowing for data gaps

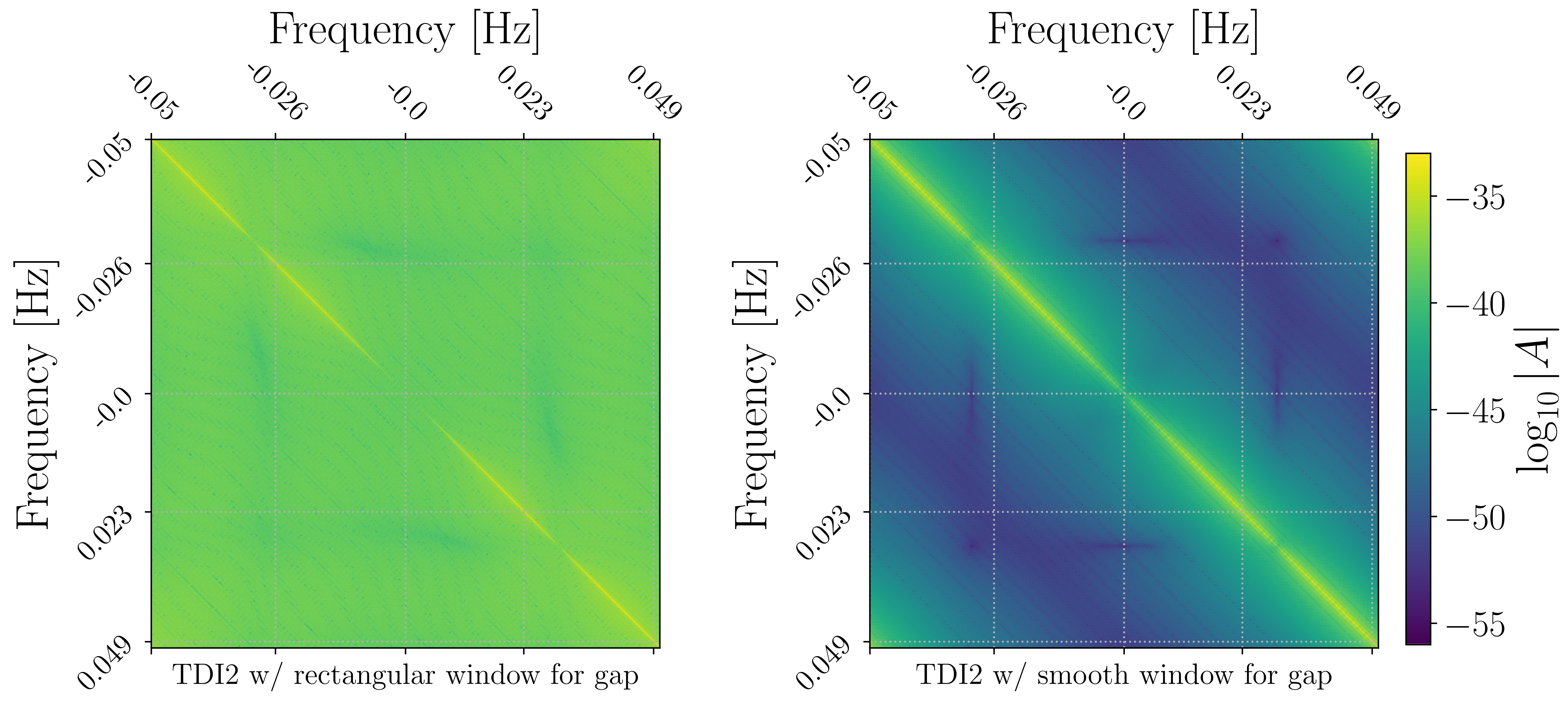
Different logics:

- Windowing only the data/signal: loss of information
- Windowing the data/signal **and** the covariance: no loss of information

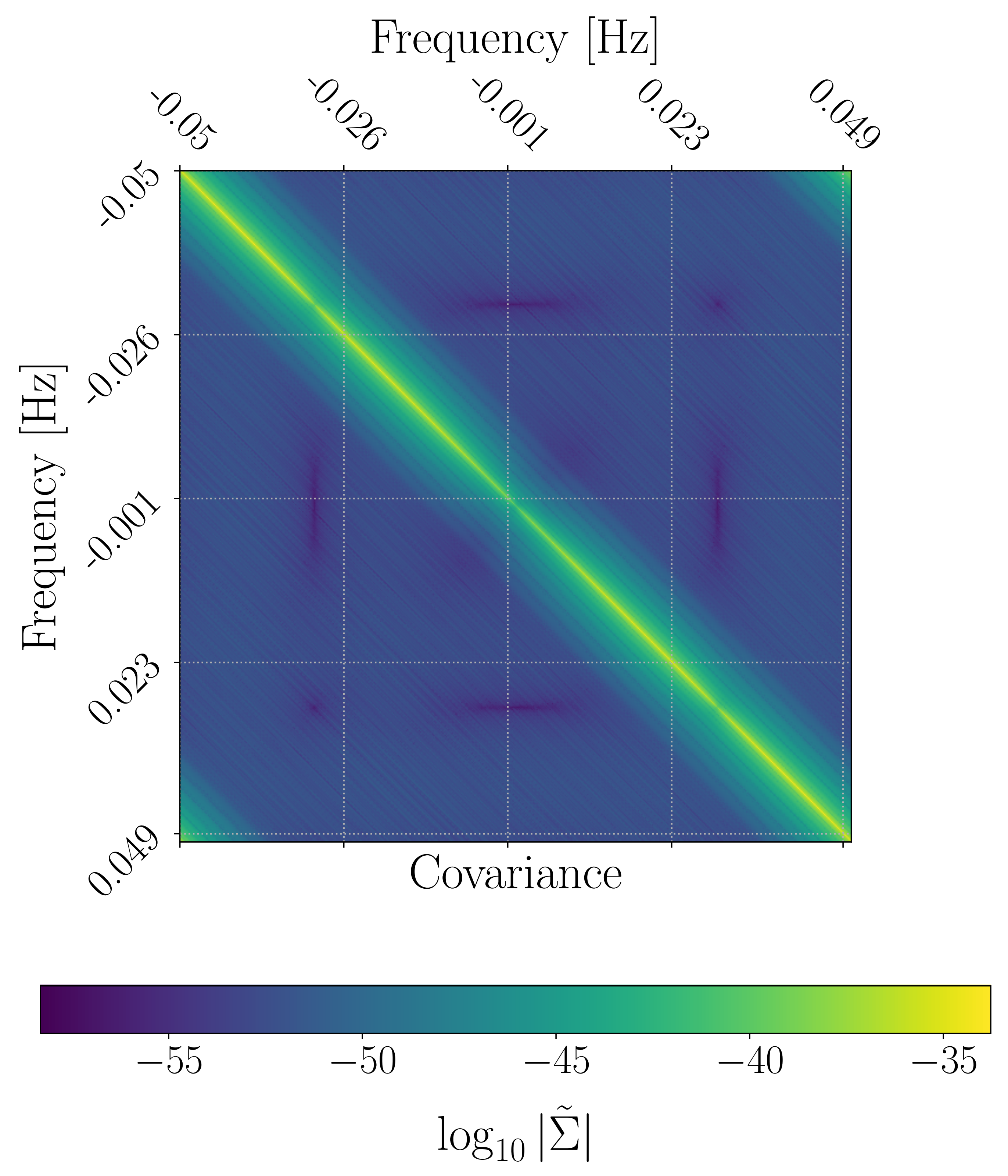
Why tapering ?

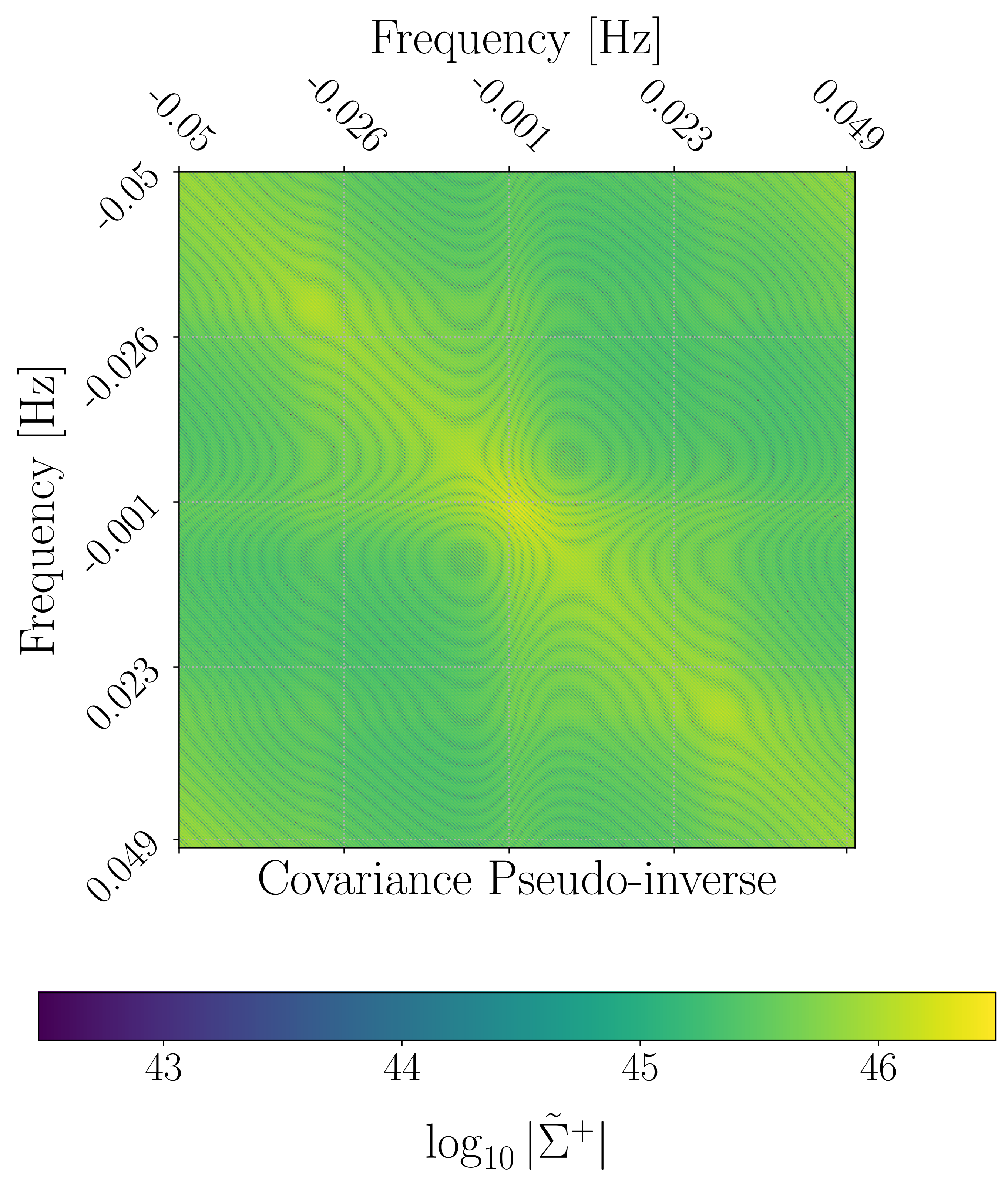
- When using the improper Whittle likelihood: tapering the gap edge, alleviate the statistical problems by avoiding a jump to 0
- When using the correct pseudo-inverse covariance: tapering preserves diagonal-dominated structure of the FD covariance, but delicate (pseudo)inversion of the matrix

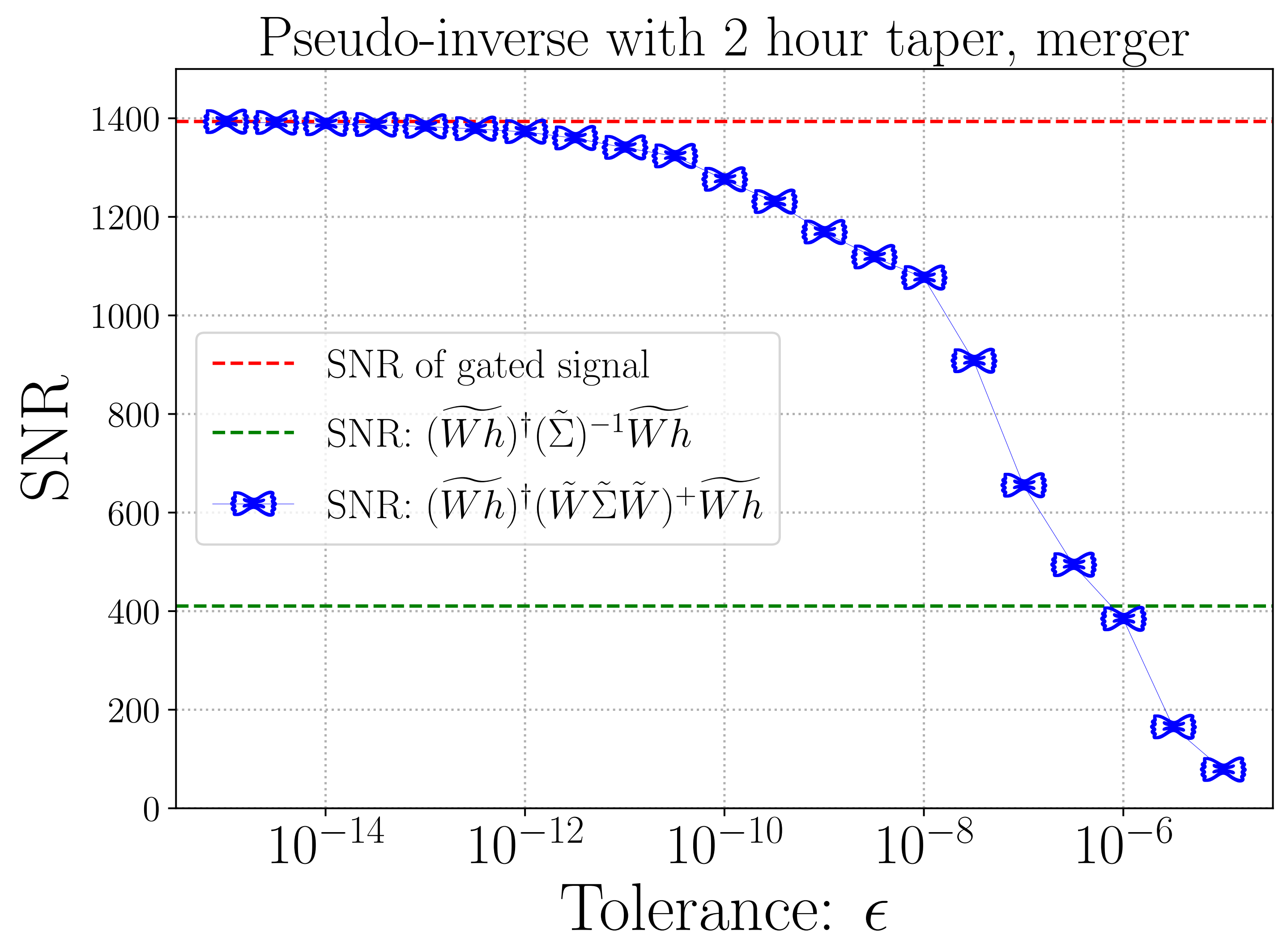




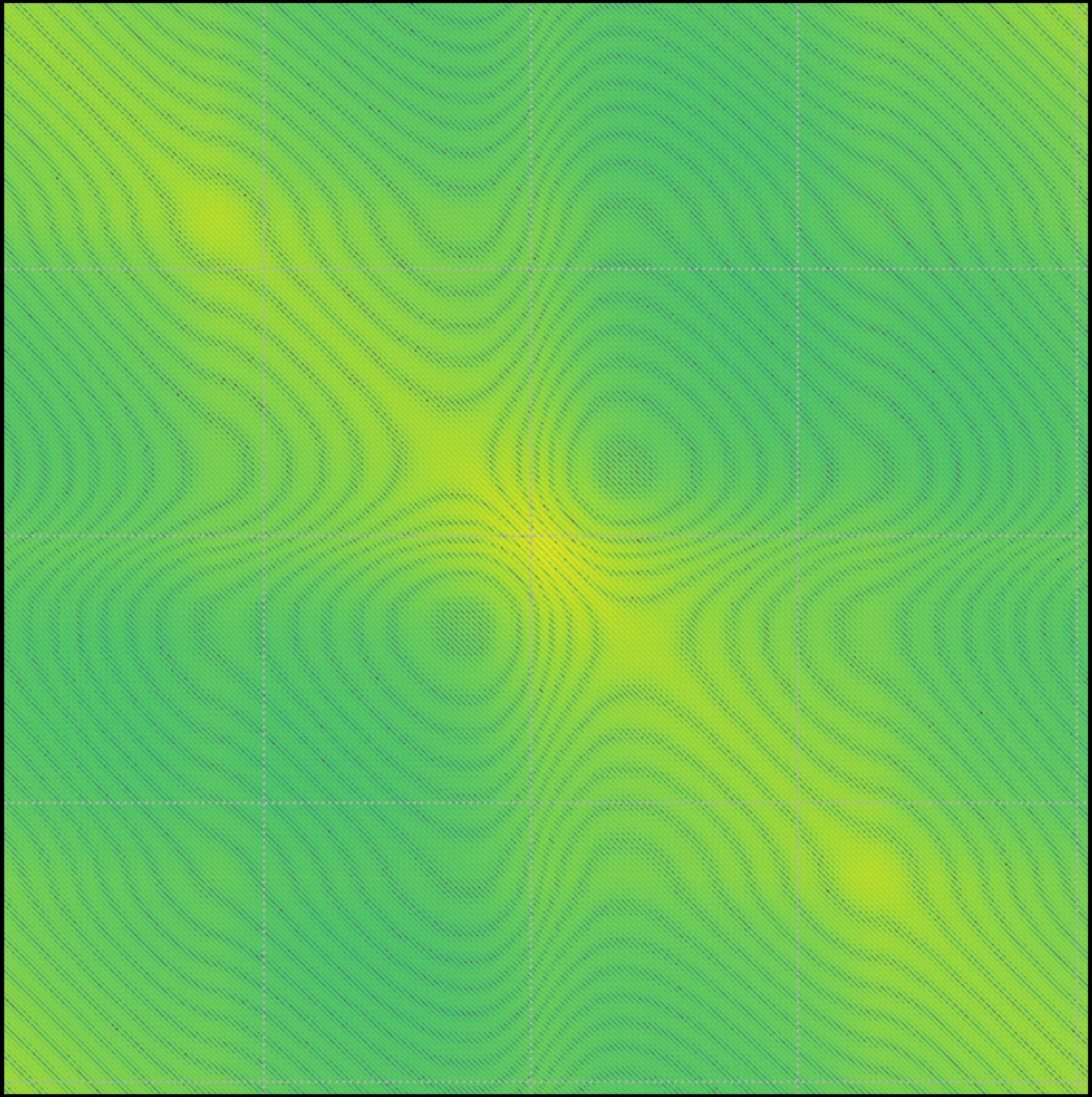
Pseudo-inverse covariance



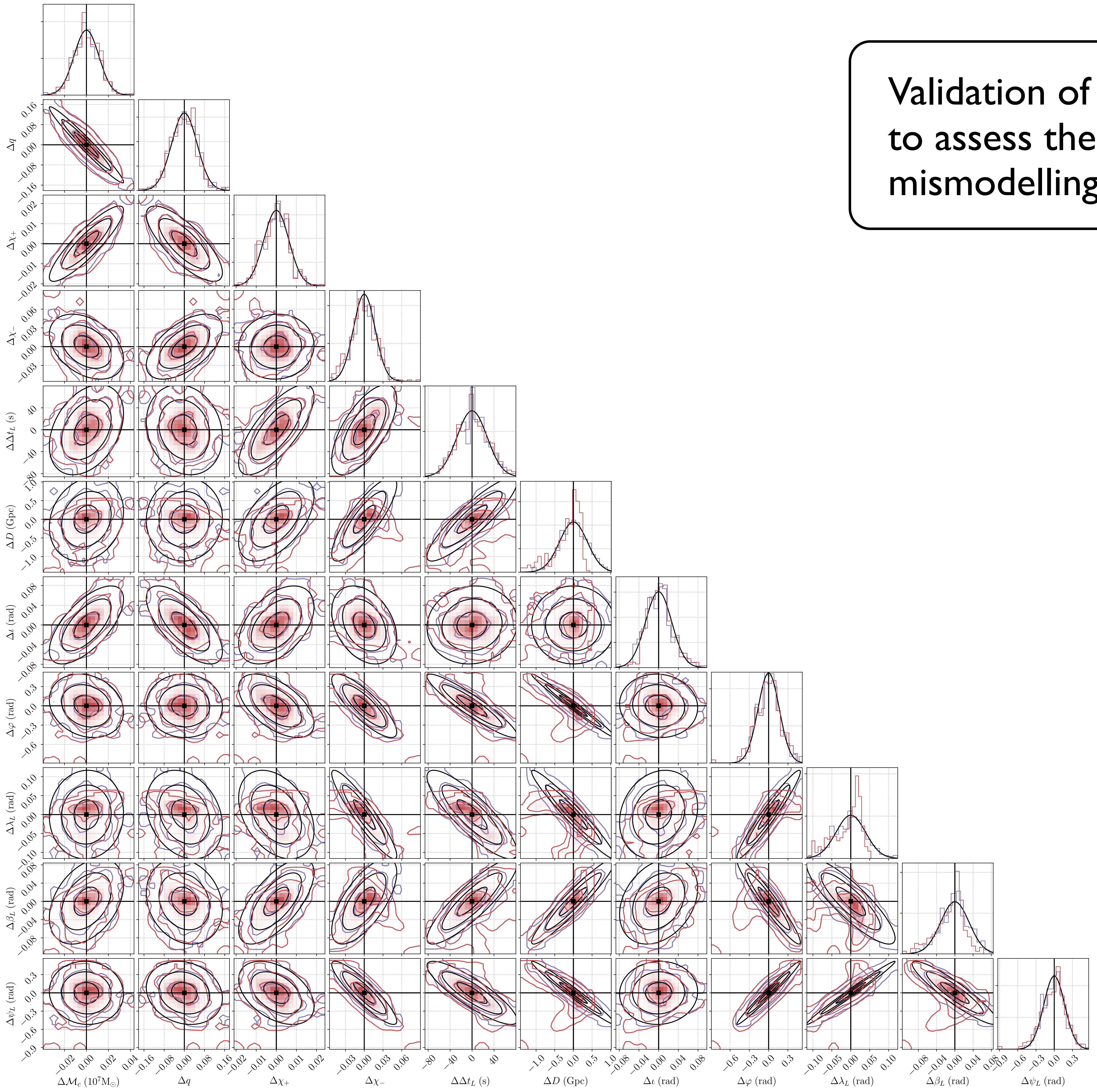




- With careful pseudo-inversion, consistent SNRs/likelihoods



Results: comparing CV mismodeling prediction to PE



Validation of the CV-inspired Υ to assess the impact of mismodelling noise